



DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITE DE COOPERATION EN MATIÈRE DE BREVETS (PCT)

(51) Classification internationale des brevets ⁶ : C12N 15/60, 15/82, A01H 5/00, C12P 21/08		A1	(11) Numéro de publication internationale: WO 95/20046 (43) Date de publication internationale: 27 juillet 1995 (27.07.95)
(21) Numéro de la demande internationale: PCT/EP95/00263 (22) Date de dépôt international: 25 janvier 1995 (25.01.95) (30) Données relatives à la priorité: 94/00787 25 janvier 1994 (25.01.94) FR (71) Déposant (pour tous les Etats désignés sauf US): BIOCEM [FR/FR]; Campus Universitaire des Cézeaux, 24, avenue des Landais, F-63170 Aubière (FR). (72) Inventeurs; et (75) Inventeurs/Déposants (US seulement): PEYRET, Pierre [FR/FR]; Rue de la Poste, F-63350 Joze (FR). ALRIC, Monique [FR/FR]; Voie Privée, 27, avenue des Landais, F-63000 Clermont-Ferrand (FR). PEREZ, Pascual [FR/FR]; 21, allée du Parc, F-63110 Beaumont (FR). (74) Mandataires: GUTMANN, Ernest etc.; Ernest Gutmann - Yves Plasseraud S.A., 3, rue Chauveau-Lagarde, F-75008 Paris (FR).			(81) Etats désignés: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LT, LU, LV, MD, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, US, UZ, VN, brevet européen (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), brevet ARIPO (KE, MW, SD, SZ). Publiée <i>Avec rapport de recherche internationale. Avant l'expiration du délai prévu pour la modification des revendications, sera republiée si de telles modifications sont reçues.</i>
(54) Title: PLANT ACONITASES AND NUCLEIC ACIDS CODING THEREFOR (54) Titre: ACONITASES VEGETALES ET ACIDES NUCLEIQUES CODANT POUR CES ACONITASES (57) Abstract <p>A protein having enzymatic aconitase activity and comprising (i) the amino acid sequence of figure 1 or a sequence at least 75 % homologous thereto, or (ii) a variant of sequence (i) comprising a deletion of up to around 40 amino acids at the NH₂ end. A plant aconitase gene and fragments derived therefrom, such as introns, coding sequences, aconitase gene promoter regions and chimeric genes including said regions, are also disclosed.</p> (57) Abrégé <p>L'invention concerne une protéine ayant une activité enzymatique d'aconitase, caractérisée en ce qu'elle comporte, soit (i) la séquence en acides aminés illustrée dans la figure (1) ou une séquence présentant au moins 75% d'homologie avec celle-ci, soit (ii) une variante de la séquence (i) comportant une délétion de jusqu'à 40 acides aminés environ au niveau de l'extrémité NH₂. L'invention concerne également le gène de l'aconitase végétale et des fragments dérivés du gène, par exemple les introns, les séquences codantes, les régions promotrices du gène de l'aconitase et des gènes chimériques comprenant ces régions.</p>			

UNIQUEMENT A TITRE D'INFORMATION

Codes utilisés pour identifier les Etats parties au PCT, sur les pages de couverture des brochures publiant des demandes internationales en vertu du PCT.

AT	Autriche	GB	Royaume-Uni	MR	Mauritanie
AU	Australie	GE	Géorgie	MW	Malawi
BB	Barbade	GN	Guinée	NE	Niger
BE	Belgique	GR	Grèce	NL	Pays-Bas
BF	Burkina Faso	HU	Hongrie	NO	Norvège
BG	Bulgarie	IE	Irlande	NZ	Nouvelle-Zélande
BJ	Bénin	IT	Italie	PL	Pologne
BR	Brésil	JP	Japon	PT	Portugal
BY	Bélarus	KE	Kenya	RO	Roumanie
CA	Canada	KG	Kirghizistan	RU	Fédération de Russie
CF	République centrafricaine	KP	République populaire démocratique de Corée	SD	Soudan
CG	Congo	KR	République de Corée	SE	Suède
CH	Suisse	KZ	Kazakhstan	SI	Slovénie
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovaquie
CM	Cameroun	LK	Sri Lanka	SN	Sénégal
CN	Chine	LU	Luxembourg	TD	Tchad
CS	Tchécoslovaquie	LV	Lettonie	TG	Togo
CZ	République tchèque	MC	Monaco	TJ	Tadjikistan
DE	Allemagne	MD	République de Moldova	TT	Trinité-et-Tobago
DK	Danemark	MG	Madagascar	UA	Ukraine
ES	Espagne	ML	Mali	US	Etats-Unis d'Amérique
FI	Finlande	MN	Mongolie	UZ	Ouzbékistan
FR	France			VN	Viet Nam
GA	Gabon				

ACONITASES VEGETALES ET ACIDES NUCLEIQUES CODANT POUR
CES ACONITASES

L'invention concerne l'aconitase végétale, des séquences d'acides nucléiques codant pour l'aconitase, ainsi que le gène de l'aconitase et des fragments dérivés de ce gène. L'invention concerne aussi les régions promotrices du gène de l'aconitase et des gènes chimériques comprenant cette région promotrice. En outre, l'invention vise un procédé pour modifier le métabolisme d'une plante par l'introduction, dans la plante, d'une séquence d'acide nucléique capable de modifier l'expression naturelle du gène de l'aconitase.

L'aconitase, (E.C.4.2.1.3.), également connu sous le nom d'aconitate hydratase ou citrate (isocitrate) hydro-lyase, est une protéine à centre fer-soufre catalysant la réaction d'interconversion réversible du citrate en isocitrate via le cis-aconitate comme intermédiaire. Depuis plusieurs années, les recherches ont conduit à la mise en évidence non seulement d'une aconitase mitochondriale (intervenant dans le cycle de Krebs) mais également d'une aconitase cytosolique, qui, chez les animaux, possède une double fonctionnalité : une activité d'aconitase, dont le rôle physiologique n'est pas connu, et une activité "Iron Responsive Element - Binding Protein" (IRE-BP) permettant une régulation post-transcriptionnelle dans le métabolisme du fer. Chez l'homme, l'IRE-BP existe dans deux états conformationnels, appelés des "isoformes", et correspondant à un état de réduction.

Chez les animaux et chez certains micro-organismes, les aconitases mitochondriales et

cytosoliques ont pu être purifiées et clonées. Par exemple, l'aconitase mitochondriale de porc (Zheng et al. 1990, J. Biol. Chem., 265 : 2814-2821) et de levure (Gangloff, 1990, PNAS, 85 : 8998-9002) a été clonée, ainsi que l'aconitase cytosolique humaine (Rouault et al., 1990, PNAS, 87 : 7958-7962) et bactérienne (Promodrou et al., 1991, J. Gen. Microbiol., 137 : 2505-2515).

Chez les animaux, l'aconitase cytosolique et l'aconitase mitochondriale sont codées par deux gènes différents et sont facilement séparables, par exemple par chromatographie sur DEAE - cellulose.

En ce qui concerne l'aconitase végétale, très peu d'informations sont disponibles à l'heure actuelle. Chez les plantes, il existe aussi une activité aconitase mitochondriale, intervenant dans le cycle de Krebs, et une activité aconitase cytosolique, intervenant dans le cycle du glyoxylate. Le fait qu'aucune activité aconitase ne soit détectée dans les glyoxysomes suggère que le cycle du glyoxylate est détournée via le cytosol (Courtois - Verniquet et Douce, 1993, Biochem. Journal, 294 : 103-107).

Un certain nombre de différences dans les propriétés structurales des aconitases mitochondriales d'origine animale et végétale, en particulier dans la structure Fer-Soufre, ont pu être mises en évidence (Brouquisse et al., 1986, Plant Physiol., 81 : 247-252). De plus, contrairement aux travaux réalisés dans le domaine animal, les formes mitochondriale et cytosolique de l'aconitase végétale n'ont pu, jusqu'à présent, être séparées par chromatographie d'échange d'ions sur DEAE (Brouquisse et al., 1987, Plant Physiol., 84 : 1402-1407), les deux formes présentant les mêmes pics d'élution. Ces résultats suggèrent que

la forme mitochondriale et la forme cytosolique de l'aconitase végétale ont des structures très voisines.

A ce jour, les aconitases de deux espèces végétales ont été purifiées. L'aconitase mitochondriale de pomme de terre a pu être purifiée à partir de mitochondries de tubercules (Verniquet et al., 1991, Biochem. J., 276 : 643-948). Il s'agit d'une protéine monomérique d'un poids moléculaire d'environ 90KDa, de nature acide, se rapprochant des aconitases bactériennes et de levure.

De Bellis et al., 1993, Physiologia Plantarum, 88 : 485-492, ont purifié deux isoformes de l'aconitase à partir de cotylédons étiolés de citrouille. Ces deux isoformes ont un poids moléculaire voisin de 100KDa et 98KDa avec des pI respectivement de 5.0 et 4.8. Selon ces auteurs, une troisième forme a également été détectée en gel non dénaturant mais n'a pas été purifiée. La localisation subcellulaire de ces isoformes n'a pas été clairement déterminée mais les auteurs supposent qu'il s'agit d'isoformes cytosoliques. La signification biologique de ces "isoformes" n'est pas élucidée.

Les informations disponibles aujourd'hui sur les aconitases végétales ne permettent donc pas de définir la relation entre la forme mitochondriale et la forme cytosolique. Il n'est pas connu si les deux formes sont codées par deux gènes différents, comme chez l'animal, ou s'il s'agit d'un seul gène. L'existence de différentes isoformes d'une même forme d'aconitase a été suggérée mais n'a pas été confirmée. De plus, l'aconitase de citrouille et celle de pomme de terre purifiées par De Bellis et par Verniquet, respectivement (supra), présentent des caractéristiques différentes, l'une de l'autre, suggérant une spécificité d'espèce. A ce jour, le degré

d'homologie entre les différentes espèces végétales n'est pas connu

Jusqu'à présent, le clonage du gène de l'aconitase végétale n'a jamais été rapporté.

La présente invention concerne le clonage de l'ADNc de l'aconitase végétale et la caractérisation du gène, y compris la région promotrice.

Le clonage du gène a présenté un certain nombre de difficultés pratiques. Par exemple, la spécificité d'espèce de l'aconitase suggérée dans l'art antérieur empêche la transposition fiable d'information concernant une première espèce, sur une deuxième. Pour cloner l'ADNc de l'aconitase de melon, les inventeurs ont donc dû purifier l'aconitase à partir de graines de melon afin d'obtenir la séquence en acides aminés de l'extrémité NH₂. Ceci a permis d'élaborer une série de sondes nucléiques dégénérées, mais aucune de ces sondes n'a permis d'identifier, parmi une banque d'ADNc, de clone positif. Une séquence partielle d'ADNc d'aconitase de melon a pu être identifiée en criblant une banque d'expression d'ADNc avec des anticorps polyclonaux, mais un clone complet de l'ADNc de l'aconitase n'a pas été identifié.

Une hybridation inter-spécifique, c'est-à-dire le criblage d'une banque d'ADNc d'Arabidopsis thaliana avec une sonde provenant de l'ADNc de l'aconitase de melon, a donc été effectuée. Les conditions d'hybridation inter-spécifique ont été spécialement conçues pour cette étape du clonage parce que l'homologie entre les séquences n'était pas connue. Un clone entier d'ADNc d'Arabidopsis thaliana a pu ainsi être identifié. Cet ADNc a ensuite servi de sonde pour cribler une banque génomique d'Arabidopsis thaliana et a permis l'identification du gène, y compris la région promotrice. A partir de ce gène, les inventeurs ont

identifié un couple d'amorces qui a pu être utilisé chez d'autres espèces notamment chez le maïs, dans une réaction d'amplification, permettant l'identification du gène de l'aconitase d'autres espèces.

Les résultats de ces travaux ont permis aux inventeurs :

- d'effectuer des recherches d'homologies de séquence entre les aconitases de différentes espèces. Il a été démontré que l'aconitase d'Arabidopsis thaliana présente, au niveau de la protéine, une forte homologie (70 %) avec l'IRE-BP humaine, et avec l'aconitase d'autres espèces végétales (90 % environ). Les protéines végétales sont plus proches des aconitases cytosoliques animales que des aconitases mitochondriales animales. Ces résultats suggèrent que l'aconitase végétale, en particulier l'aconitase cytosolique, joue le rôle d'une IRE-BP et permet une régulation post-transcriptionnelle dans le métabolisme d'ions métalliques tels que le fer, ou bien qu'elle joue le rôle d'une RNA-BP, se fixant à une structure proche ou similaire des IRE animales, qui interviendrait dans une régulation post-transcriptionnelle du métabolisme de la plante ou de la graine ;

- de caractériser le gène de l'aconitase chez Arabidopsis thaliana. Sa structure est très complexe, comportant 20 exons et 19 introns. Cinq départs de transcription ont été cartographiés, dont un est utilisé préférentiellement. Un fragment présentant une activité promotrice a été identifié. Ce promoteur permet une expression constitutive particulièrement forte lors de la germination des graines et de la maturation des grains de pollen ;

- de mettre en évidence l'existence d'un clone putatif dont l'initiation de transcription se

situerait en amont de la TATA box identifiée. Ce gène pourrait donc coder pour les deux formes protéiques, son expression étant éventuellement dirigée par deux promoteurs ;

- d'émettre l'hypothèse que la forme mitochondriale de l'aconitase se différencie de la forme cytosolique essentiellement par la présence, à l'extrémité NH_2 , d'un peptide signal mitochondriale.

- de caractériser le gène de l'aconitase chez le maïs et ainsi, de faire une comparaison entre l'aconitase de plantes dicotylédones et celle de plantes monocotylédones. La protéine est extrêmement conservée (90% d'homologie) entre les deux types de plantes, et présente des régions spécifiques aux plantes, et d'autres conservées, parmi toutes les espèces étudiées, y compris les animaux et les micro-organismes. Au niveau du gène, l'organisation génomique est identique chez les monocotylédones et chez les dicotylédones.

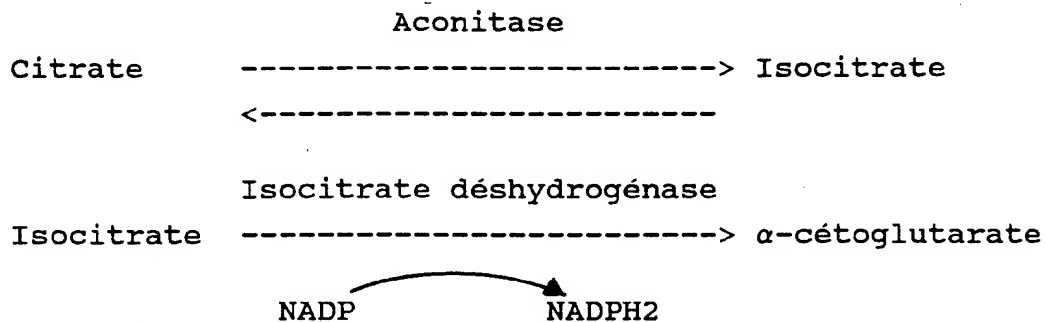
L'objet de la présente invention est une protéine ayant une activité enzymatique d'aconitase. Plus particulièrement l'invention concerne une protéine ayant une activité enzymatique d'aconitase caractérisée en ce qu'elle comporte, soit

- i) la séquence en acides aminés illustrée dans la figure 1 ou une séquence présentant au moins 75 % d'homologie avec celle-ci, soit :

- ii) une variante de la séquence i) comportant une délétion jusqu'à 40 acides aminés environ au niveau de l'extrémité NH_2 .

Dans le contexte de la présente invention, une activité enzymatique d'aconitase est définie comme étant la capacité de catalyser la réaction d'isomérisation du citrate en iso-citrate via le cis-aconitate. Cette activité peut être dosée par la

méthode décrite par Rose et O'Connel (J. Biol. Chem., 242 : 1870-1879, 1967). Cette méthode fait appel au couplage de plusieurs réactions enzymatiques décrites ci-dessous :



La réduction du NADP est suivie par spectrophotométrie à 340nm.

Une autre méthode de dosage d'activité enzymatique de l'aconitase est celle décrite par Kennedy et al. (J. Biol. Chem., 258 : 11098-11105, 1983) qui permet de suivre à 240 nm l'apparition du cis-aconitate à partir de l'isocitrate. La réaction enzymatique se déroule à 25° C dans un tampon phosphate 50mM à pH 6 avec isocitrate 50mM comme substrat. Le coefficient d'extinction molaire du cis-aconitate est donné comme étant égal à 3,6mM⁻¹, une unité enzymatique (U) correspondant à la formation d'une micro-mole de cis-aconitate par minute. Ce dosage, facile à mettre en oeuvre, est utilisé pour déterminer les constantes enzymatiques de l'aconitase et pour suivre sa purification.

La protéine de l'invention, outre son activité d'aconitase, peut également présenter une activité d'IRE-BP, c'est-à-dire, est capable de se fixer à des ARNm ayant une structure secondaire voisine de celle de l'iron-responsive element (IRE) animal.

La protéine de l'invention, ayant une activité enzymatique d'aconitase, peut être celle dont la

séquence en acides aminés est illustrée dans la figure 1, et qui représente l'aconitase d'*Arabidopsis thaliana*.

L'invention englobe également des protéines présentant au moins 75 % d'homologie avec la séquence illustrée dans la figure 1 et ayant une activité d'aconitase. Dans le contexte de la présente invention, le pourcentage d'homologie entre deux séquences d'acides aminés est calculé comme étant le nombre d'acides aminés identiques plus le nombre d'acides aminés similaires dans l'alignement des deux séquences, divisé par la longueur des séquences entre deux positions données. Si, entre les deux positions données, les deux séquences n'ont pas la même longueur, le pourcentage d'homologie est le nombre d'acides aminés identiques et similaires, divisé par la longueur de la séquence la plus longue. Les acides aminés "similaires" sont connus dans l'art, voir par exemple R.F. Feng, M.S. Jobson and R.F. Doolittle, 1985, J. Mol. Evol., 21 : 112-115. Ils sont normalement considérés comme étant ceux qui, au sein d'une matrice de permutation, ont un coefficient de substitution positif.

De préférence, les protéines homologues de l'invention présentent au moins 80, et plus particulièrement au moins 85, par exemple 90 ou 95 % d'homologie, et de préférence au moins 80% par exemple 85% ou 90% d'identité, avec celle illustrée dans la figure 1.

Comme protéines préférées de l'invention, on peut aussi citer les deux aconitases de melon purifiées par les inventeurs à partir de graines de melon. Les deux formes semblent avoir en commun une même séquence en acides aminés. La séquence partielle en acides aminés est illustrée dans la figure 2. Elle présente

86,5 % d'identité avec celle d'Arabidopsis thaliana (91,8 % d'homologie). Les deux formes d'aconitase de melon purifiées par les inventeurs, selon la méthode décrite dans les exemples plus loin, ont un poids moléculaire apparent estimé en gel de polyacrylamide en conditions dénaturantes voisin de 97KDa. Les points isoélectriques de ces deux protéines (appelées "AcoI" et "AcoII") sont estimés respectivement à 5,2 et 5.0.

Comme autre exemple de protéine préférée de l'invention, on peut citer l'aconitase de maïs (figures 18 et 20)

Les protéines présentant au moins 75 % d'homologie avec celle de la figure 1 comprennent aussi l'aconitase d'autres espèces végétales, en particulier les Angiospermes. On peut citer en particulier l'aconitase des dicotylédones comprenant les légumineuses, les crucifères, les solanacées, les cucurbitacées, les chénopodiacées, les ombellifères, le tournesol, le soja. Les cucurbitacées et les espèces du genre Brassica, par exemple B. napus, B. oleracea etc.. sont particulièrement préférées. On peut aussi citer l'aconitase des monocotylédones comprenant les céréales telles que le blé, le maïs, l'orge, le triticale et le riz.

Les protéines de l'invention peuvent également être définies comme étant des protéines ou des enchaînements protéiques comportant au moins l'une des séquences protéiques suivantes, ou une séquence présentant au moins 60%, et de préférence au moins 70% ou 80% d'homologie avec l'une de ces séquences, ces séquences étant spécifiques aux aconitases végétales :

I : DTVSMIEAYLRANKMFVDYSEPEskTVYSSCLELNLEDVE

II : PLKEMKADWHSCLDNRV

10

III : AVPKEAQS KAVEFNFN GTTAQLR

IV : KGMTMSPPG

V : AVVMMRLWREHFANIRIVNKHLKGEVG

VI : NVSEIKPGQDVTVTNN

VII : les acides aminés 1 à 40 de l'extrémité NH₂ de la protéine illustrée dans la figure 1.

Ces séquences se trouvent aux positions suivantes de l'aconitase d'Arabidopsis thaliana illustrée dans la figure 2 :

I : 361-400

II : 414-430

III : 436-458

IV : 683-691

V : 745-771

VI : 875-891

Outre les régions spécifiques aux plantes, les protéines de l'invention renferment également plusieurs zones hautement conservés au sein de toutes les espèces étudiées. Ces zones seront appelées dans ce qui suit "régions conservées du point de vue phylogénétique". Il s'agit de séquences d'acides aminés ayant une longueur d'au moins 7 acides aminés et étant conservées avec au moins 75% d'homologie (environ 60% d'identité), et souvent au moins 80% d'homologie, chez les plantes, les animaux et les micro-organismes. Comme exemple de ce type de région, on peut citer les séquences suivantes (les numéros entre parenthèses étant les acides aminés d'A. thaliana, numérotés selon la figure 2) :

VIII : PGSGIVHQVNLE (acides aminés 204 à 216),

IX : GTDSHTTMIDGLG (acides aminés 236 à 248)

X : DSITTDHISPAG (acides aminés 406 à 417)

XI : GSGSSRD (acides aminés 808 à 814)

L'invention englobe également des variantes de la séquence illustrée dans la figure 1 et de celles présentant au moins 75 % d'homologie avec celle-ci, les variantes comportant une délétion de jusqu'à 40 acides aminés environ au niveau de l'extrémité NH₂. Cette délétion n'affecte pas l'activité enzymatique d'aconitase de la protéine. Du point de vue physiologique, cette délétion correspond soit à l'élimination du peptide signal mitochondrial soit à l'élimination de la partie de la séquence qui différencie la forme mitochondriale de la forme cytosolique. De préférence, la délétion s'étend de l'acide aminé 1 à l'acide aminé 30 environ.

Pour Arabidopsis thaliana, la délétion au niveau de l'extrémité NH₂ s'étend de préférence de l'acide aminé 1 à l'acide aminé 27 (numérotation selon la figure 1) ou encore de l'acide aminé 1 à l'acide aminé 29. En effet, les inventeurs démontrent que la forme mitochondriale de l'aconitase d'Arabidopsis thaliana démarre à l'acide aminé 28 (Ser Ser Met...). Ceci signifie que les acides aminés 1 à 27 correspondent vraisemblablement au peptide signal mitochondrial. La forme cytosolique de l'aconitase d'Arabidopsis thaliana démarre vraisemblablement à l'acide aminé 30 (Met).

Dans ce qui suit, les protéines ayant une activité d'aconitase et comportant la séquence de la figure 1, ou une séquence présentant au moins 70 % ou 75% d'homologie avec celle-ci, ou encore les variantes comportant une délétion NH₂-terminale, seront appelées "les protéines de l'invention".

L'invention concerne également des fragments protéiques, ou des "enchaînements d'acides aminés", comprenant au moins 6 acides aminés de la protéine de l'invention, et de préférence au moins 8 ou au moins 10 acides aminés. La longueur maximale des fragments de l'invention correspond à la longueur de la protéine entière moins un acide aminé. Typiquement, les fragments de l'invention ont une longueur comprise entre 10 et 900 acides aminés, par exemple entre 20 et 800, ou entre 30 et 500 acides aminés, notamment 50, 200, 300, 400, etc... acides aminés. Les fragments protéiques de longueur importante, par exemple au-delà de 500 acides aminés, peuvent éventuellement présenter une activité enzymatique d'aconitase. Les fragments peuvent donc être des peptides, des polypeptides ou des protéines. Ces fragments protéiques peuvent être reconnus par des anticorps capables de reconnaître la protéine entière, les fragments présentant donc une homologie immunologique avec la protéine "mère".

Selon une variante de l'invention, les fragments ont une longueur comprise entre 6 et 40 acides aminés et présentent, sur toute leur longueur, une homologie d'au moins 90%, et de préférence au moins 90% d'identité, avec la partie correspondante de la protéine de la figure 1. L'expression "la partie correspondante" signifie la partie de la protéine illustrée dans la figure 1 avec laquelle l'alignement est fait pour calculer l'homologie. Ces fragments ont typiquement une longueur comprise entre 8 et 30 acides aminés.

Exceptionnellement, les fragments de l'invention peuvent avoir une longueur inférieure à 40 acides aminés et ne présenter que 60% d'homologie avec la partie correspondante de la protéine de la figure 1. Ceci est le cas lorsqu'il s'agit de fragments

correspondant aux régions de la protéine spécifiques aux plantes. Dans le contexte de l'invention, "spécifique aux plantes" signifie une séquence d'acides aminés ayant une longueur d'au moins 6 et de préférence au moins 8 acides aminés, présentant moins de 40%, et souvent moins de 30% d'homologie avec les séquences correspondantes chez les mammifères et chez les organismes unicellulaires. Au sein du règne végétal, ces régions sont conservées avec une homologie d'au moins 60%, typiquement au moins 80%. Les régions spécifiques aux aconitases végétales comprennent les séquences suivantes :

I : DTVSMIEAYLRANKMFVDYSEPEKSTVYSSCLELNLEDVE

II : PLKEMKADWHSCLDNRV

III : AVPKEAQSKAVEFNFNNGTTAQLR

IV : KGMTMSPPG

V : AVVMMRLWREHFANIRIVNKHLLKGEVG

VI : NVSEIKPGQDVTVTNN

dont les emplacements ont été indiqués plus haut.

Selon ce dernier mode de réalisation, le fragment peut être constitué d'un fragment de l'extrémité NH₂ de la protéine illustrée dans la figure 1, l'expression "extrémité NH₂", signifiant la partie de la protéine s'étendant de l'acide aminé 1 à l'acide aminé 40, par exemple entre 1 et 30. Le fragment peut également être constitué d'un sous-fragment d'une séquence ayant une longueur de 40 acides aminés environ et présentant au moins 60 % d'homologie avec l'extrémité NH₂ de la protéine illustrée dans la figure 1. Le fragment NH₂-terminal a normalement une longueur comprise entre 6 et 40 acides aminés, par

exemple 30. De préférence, ces fragments présentent au moins 70 %, et plus particulièrement au moins 80 % d'homologie, avec les acides aminés 1 à 40 de celle illustrée dans la figure 1. L'extrémité NH₂ correspond vraisemblablement au peptide signal mitochondrial responsable du transport de l'aconitase dans les mitochondries. Les fragments selon cet aspect de l'invention constituent donc des peptides signaux et peuvent être utilisés, par exemple dans la production de protéines de fusion pour assurer le transport dans les mitochondries.

Selon une autre variante de l'invention, les fragments protéiques ont une longueur supérieure à 40 acides aminés, par exemple 40 à 100, et présentent sur toute leur longueur au moins 75%, et de préférence au moins 80 ou 90% d'homologie avec la partie correspondante de la protéine de la figure 1. Comme exemples de ce type de fragment, on peut citer la moitié NH₂ terminale de la protéine, la moitié COOH terminale ou encore la partie centrale.

Les enchaînements d'acides nucléiques de l'invention peuvent comprendre, outre le fragment dérivé de l'aconitase végétale décrite ci-dessus, d'autres régions ne présentant aucune relation avec l'aconitase. Cette variante englobe des protéines de fusion et d'autres protéines "hybrides".

Les protéines et fragments protéiques de l'invention peuvent être obtenus par purification à partir de la plante, en mettant en oeuvre le protocole expérimental décrit dans les exemples ci-dessous, suivi éventuellement par un clivage de la protéine pour obtenir des fragments. Ils peuvent également être produits par voie recombinante après l'insertion dans un hôte cellulaire compétent d'une séquence d'acide nucléique appropriée suivie de la purification de la

protéine ainsi produite. Ils peuvent également être obtenus par synthèse chimique.

Outre les protéines et fragments protéiques, l'invention concerne également les gènes de l'aconitase végétale, et toutes les séquences d'acide nucléique dérivées de la séquence génomique, toutes les séquences codantes y compris les séquences dégénérées, les séquences régulatrices, les introns, les séquences susceptibles de s'hybrider avec le gène de l'aconitase végétale et des fragments de ces séquences.

Plus particulièrement, l'invention concerne une séquence d'acide nucléique caractérisée en ce qu'elle comporte soit :

i) une séquence qui, le cas échéant après transcription, épissage et traduction, donne lieu à la protéine de l'invention, soit

ii) le transcrit de la séquence i) ou son équivalent ADNc, soit

iii) une séquence complémentaire de la séquence i) ou ii), soit

iv) une séquence ayant une longueur d'au moins 25 bases et capable de s'hybrider, sur toute sa longueur, avec les séquences i), ii) ou iii) dans des conditions stringentes ou moyennement stringentes, soit

v) un fragment consistant en au moins 6, et de préférence au moins 25 ou 30, nucléotides consécutives de l'une quelconque des séquences i), ii), ou iii).

Le terme acide nucléique signifie dans le cadre de la présente invention de l'ADN ou de l'ARN, ou éventuellement un mélange des deux.

Selon cet aspect de l'invention, une séquence d'acides nucléiques du type i) est la séquence génomique d'Arabidopsis thaliana illustrée dans la

figure 1 ou celle du maïs illustrée dans la figure 18. On peut aussi citer les séquences génomiques végétales du gène de l'aconitase chez d'autres espèces, par exemple les monocotylédones, telles que les céréales etc..., et dicotylédones, par exemple de nombreux légumes. Ces séquences peuvent être identifiées par l'utilisation d'une réaction d'amplification d'acide nucléique, de préférence Reverse-Transcriptase Polymerase Chain Reaction (RT-PCR), les amorces et les conditions étant, par exemple, celles utilisées dans l'exemple 11 plus loin, le substrat étant de l'ARN poly A+. On peut également utiliser d'autres amorces dérivées du gène de l'aconitase végétale, par exemple les séquences codant pour tout ou partie des enchaînements d'acides aminés I à XI cités ci-dessus.

Les séquences génomiques peuvent également être identifiées chez des espèces autres que celles spécifiquement exemplifiées dans cette demande par l'utilisation d'une sonde spécifique au gène de l'aconitase végétale, par exemple une sonde contenant les séquences I à VI citées ci-dessus. Les conditions d'hybridisation sont celles citées dans l'exemple 6.

Comme exemple de séquences d'acide nucléique du type ii), on peut citer les transcrits (mRNA) du gène de l'aconitase végétale, ou son équivalent en ADNc, par exemple les ADNc illustrés dans les figures 3 et 4, ainsi que toutes les séquences dégénérées capables de coder pour la protéine de la revendication 1. Les séquences du type ii) englobent également le transcrit primaire (pré-mRNA) c'est-à-dire le transcrit de l'ADN génomique avant épissage. Le pré-mRNA contient notamment les introns.

L'invention englobe également les séquences du type iii), c'est-à-dire les séquences complémentaires des séquences i) et ii). L'expression "complémentaire"

signifie normalement une complémentarité de 100 %, mais peut également signifier un degré de complémentarité suffisamment élevé pour permettre l'hybridation stable entre la séquence et sa séquence complémentaire dans des conditions stringentes. La présence d'un certain nombre de mésappariements, par exemple jusqu'à 10 %, ou éventuellement jusqu'à 20 %, peut être tolérée.

L'invention concerne également des séquences du type iv), c'est-à-dire des séquences d'acides nucléiques ayant au moins 25 bases et étant capables de s'hybrider avec les séquences i), ii) et iii) telles que définies ci-dessus, dans des conditions stringentes ou moyennement stringentes. De préférence, ces séquences ont une longueur entre 25 et 1000 bases, par exemple 30 à 900 ou 50 à 500.

Les conditions stringentes sont les suivantes :

- l'hybridation s'effectue sur les filtres de nitro-cellulose dans 20ml de solution d'hybridation (SSC 6X, SDS 5%, DENHARDT 5X, ADN de spermes de saumon 50mg/ml).

- une pré-hybridation de 4 heures est suivie d'une hybridation de 12 heures (solution contenant la sonde marquée) à la température de 68°C.

Les filtres subissent les lavages suivants :

- 2 lavages de 30 mn dans une solution SSC 2X à 68°C ;

- un lavage de 30 mn dans une solution SSC 2X, SDS 0.1% à 68°C ;

- un lavage de 30 mn dans une solution SSC 0.1X, SDS 0.1% à 68°C.

Les conditions de moyenne stringence (à appliquer par exemple lors du criblage d'une banque provenant d'une espèce végétale avec une sonde provenant d'une autre espèce végétale) sont les suivantes :

- les conditions d'hybridation se font en faisant varier la concentration de formamide 10, 20, 30, 40, 50 % à une température fixe de 40°C dans un tampon d'hybridation SSC 6X, SDS 0.1%, DENHARDT 5X, ADN de spermes de saumon 50mg/ml.

- une pré-hybridation de 4 heures est suivie d'une hybridation de 12 heures.

Les conditions de lavage sont les suivantes :

- un lavage de 30 mn dans une solution SSC 2X à 42°C ;

- un lavage de 30 mn dans une solution SSC 1X à 42°C ;

- un lavage de 30 mn dans une solution SSC 0.1X à 42°C ;

- un lavage de 30 mn dans une solution SSC 0.1X, SDS 0.1% à 68°C.

Ces conditions de stringence permettent d'identifier les séquences codant pour l'aconitase chez une espèce végétale par le criblage d'une banque d'ADNc de cette espèce avec une sonde provenant d'une autre espèce végétale.

En général, les conditions stringentes permettent l'hybridation de deux séquences d'acide nucléique présentant au moins 85%-90% d'homologie. Les conditions moyennement stringentes permettent l'hybridation de séquences présentant au moins 60% ou au moins 65% d'homologie. Dans le contexte de l'invention, "homologie" au niveau de l'acide nucléique signifie identité.

L'invention vise également des séquences du type (v), c'est-à-dire des fragments d'acide nucléique consistant en au moins 6, et de préférence au moins 20, 25 ou 30, par exemple 40, 50, 60 ou plus nucléotides consécutives de l'une quelconque des séquences (i), (ii), ou (iii) précédemment décrites.

Avantageusement, le fragment comporte entre 10 et 500 nucléotides, par exemple entre 30 et 300.

Le fragment peut être capable de jouer le rôle d'amorce dans une réaction d'amplification d'acide nucléique. Dans ce cas, un couple de fragments est utilisé, chaque séquence ayant une longueur de 15 à 300 bases environ, de préférence entre 18 et 50 bases. Les fragments choisis comme amorces doivent avantageusement correspondre à des zones de la protéine séparées les unes des autres par moins de 900 acides aminés, de préférence moins de 500, par exemple une cinquantaine d'acides aminés.

Le fragment dérivé du gène de l'aconitase peut faire partie d'une séquence "hybride" plus grande, comportant d'une part au moins un fragment selon l'invention et d'autre part, une ou plusieurs séquences ne présentant pas d'apparenté avec le gène de l'aconitase, par exemple une séquence dérivée d'un deuxième gène.

Le fragment peut être marqué par des moyens connus en soi, par exemple des moyens radioactifs, fluorescents, enzymatiques, ou par un marquage à l'avidine ou à la biotine. Selon cet aspect de l'invention, le fragment marqué sert de sonde nucléotidique. Dans ce cas, la longueur du fragment est de préférence comprise entre 6 et 900 bases, par exemple entre 50 et 350 bases.

Selon une variante préférée de l'invention, la séquence d'acide nucléique est une séquence dite "complémentaire" qui comporte un ou plusieurs fragments ayant une longueur d'au moins 6 nucléotides, par exemple au moins 10 nucléotides, le(s)dit(s) fragment(s) étant complémentaire(s) d'une partie, au moins, du transcrit du gène de l'aconitase. Ce genre de structure, capable d'inhiber l'expression du gène

de l'aconitase, englobe plus particulièrement des séquences antisens et des ribozymes dirigés contre le transcrit du gène de l'aconitase.

Dans le cas de l'antisens, la séquence comporte normalement un seul fragment, complémentaire d'une partie au moins, du transcrit, ce fragment ayant une longueur d'au moins une vingtaine de nucléotides. Plusieurs séquences antisens peuvent être reliées les unes aux autres soit de manière contigüe, soit en alternance avec d'autres séquences.

De préférence, la séquence inhibitrice est un ribozyme dirigé contre le transcrit du gène de l'aconitase, le ribozyme comprenant une séquence complémentaire d'une partie au moins du transcrit du gène de l'aconitase et, incluse au sein de cette séquence complémentaire, une région catalytique provenant d'un ribozyme, de préférence un ribozyme du type tête de marteau (voir par exemple EP-A-321 201). Le ribozyme peut être un mono-ribozyme ou un poly-ribozyme.

Selon l'invention, les séquences antisens et les ribozymes sont normalement produits in-vivo par la transcription de la séquence d'ADN correspondante, insérée de manière stable dans le génome de la plante. La traduction du transcrit du gène de l'aconitase est alors empêchée, soit par l'hybridation de la séquence antisens, soit par le clivage du transcrit par le ribozyme, affectant ainsi le fonctionnement du cycle de Krebs et du cycle glyoxylate, ainsi que la ou les voie(s) métabolique(s) régulée(s) post-transcriptionnellement par l'aconitase cytosolique grâce à son activité RNA-binding. L'inhibition de la production d'aconitase conduit à une réduction du taux de production des acides organiques de ces deux cycles, mais à une augmentation de la production de

l'acétyl coenzyme-A. Ce dernier est le produit de la dégradation des squelettes carbonés de plusieurs voies du catabolisme, par exemple la glycolyse, le cycle des pentoses, la β -oxydation des lipides et le catabolisme des protéines. La dégradation de ces substances peut être donc indirectement modifiée par l'inhibition de l'expression de l'aconitase.

La mise en oeuvre et les applications de la séquence inhibitrice est décrite en détail plus loin.

Selon un mode de réalisation préférée de l'invention, le fragment d'acide nucléique est un fragment de la séquence génomique du gène de l'aconitase végétale, par exemple celle illustrée dans la figure 1 ou celle illustrée dans la figure 18. Le fragment génomique comporte au moins 25 nucléotides et de préférence au moins 30, par exemple entre 50 et 500 ou plus, sauf s'il s'agit d'une partie d'une région non codante ou d'un intron de la séquence génomique, auquel cas, le fragment peut avoir une longueur d'au moins 10 nucléotides.

Les introns du gène de l'aconitase végétale constituent des fragments préférés de l'invention. Il a été constaté que l'organisation génomique du gène de l'aconitase est identique chez Arabidopsis (dicotylédone) et chez le maïs (monocotylédone), les introns se positionnant exactement de la même manière.

Chez les monocotylédones et chez les dicotylédones, le gène de l'aconitase contient un nombre d'introns particulièrement élevé : environ 19 à 20. L'aconitase représente donc une source d'introns très riche. Par ailleurs, l'épissage des introns doit être particulièrement efficace dans la mesure où le gène est exprimé de manière constitutive.

L'invention englobe donc les introns des gènes d'aconitase végétale, y compris les signaux

d'épissage, et des fragments des introns comprenant au moins 10, et de préférence au moins 30 bases. En général, la longueur des introns varie entre 70 et 500 bases environ. Selon cette variante, les séquences comprenant le ou les introns peuvent également comprendre les séquences qui, dans le génome, jouxtent immédiatement l'intron. Le fragment peut alors comprendre de part et d'autre de l'intron, une séquence correspondant à une partie de l'exon qui naturellement jouxte l'intron, cette séquence ayant de préférence une longueur comprise entre 1 et 100 bases, par exemple 5 à 50.

Comme exemple d'introns de l'invention, on peut citer ceux d'Arabidopsis et de maïs (voir tableau III, exemple 12). On peut également citer les introns d'autres espèces, les séquences génomiques pouvant être identifiées de la manière indiquée dans l'exemple 12.

Les introns du gène de l'aconitase végétale peuvent être utilisés comme éléments susceptibles d'augmenter l'efficacité d'expression d'une séquence hétérologue chez une plante. En effet, il a été démontré, particulièrement chez les monocotylédones que l'insertion d'un intron dans la partie 5' non traduite d'un gène, c'est-à-dire entre le site d'initiation de transcription et le site d'initiation de traduction, conduit à une amélioration de la stabilité du messager, et par conséquent, à une meilleure expression. Le ou les introns utilisés de cette manière proviennent de préférence d'une monocotylédone telle que le maïs. Il s'agit de préférence, mais pas obligatoirement, du premier intron du gène. Le premier intron du gène de l'aconitase de maïs peut être obtenu par séquençage de la partie du gène qui se trouve en amont de la

séquence génomique partielle de la figure 18 contenue dans le clone R1 (voir exemple 12). Le séquençage complet du fragment SalI-HindIII de 6800 pb (HindIII = position 2841) devrait permettre l'obtention du premier intron et du promoteur (voir exemple 12).

Comme autre exemple de fragment d'acide nucléique non-codant, on peut citer le fragment -85 à +1, par exemple -85 à -52, de la séquence de la figure 1. Il a été constaté que cette région présente une structure secondaire typique des IRE (Iron Responsive Element). Il se pourrait que cette partie de la séquence soit spécifique du messenger de l'aconitase mitochondriale, auquel cas un ribozyme dirigé spécifiquement contre cette région, ou une partie, par exemple la région -49 à +1, permettrait d'inhiber spécifiquement l'activité mitochondriale.

Un autre fragment particulièrement préféré est le fragment HindIII - BglIII (-1293 à +45 de la séquence génomique de la figure 1), ou le fragment correspondant dans des séquences codant pour des protéines homologues. Le promoteur du gène de l'aconitase se trouve au sein de ce fragment. Ce fragment peut être utilisé en tant que tel comme promoteur pour obtenir l'expression de séquences codantes hétérologues chez les plantes.

Des fragments de plus petite taille peuvent également être utilisés à condition que l'activité promotrice soit conservée. Ces "sous-fragments" peuvent être identifiés par la fusion du sous-fragment en question avec la séquence codante d'un gène rapporteur tel que celui du beta glucuronidase, et la détection de l'expression du gène rapporteur dans des cellules végétales. La délétion, au sein du fragment Hind III - Bgl II, de 100 à 300 nucléotides du côté 5' donne lieu normalement à des sous-fragments ayant

une activité promotrice. Il est préférable de préserver la phase de lecture pour tester le fragment choisi.

Il est également possible, selon l'invention, d'utiliser deux régions promotrices en tandem pour augmenter le niveau d'expression.

Le promoteur conduit à une expression constitutive et est fortement exprimé à la fois lors de la germination des graines et lors de la maturation de la graine et des grains de pollen. Il est donc particulièrement avantageux d'utiliser la région promotrice de l'invention pour obtenir une forte expression pendant cette période de la vie de la plante. Les tissus de la graine, où l'expression dirigée par le promoteur est particulièrement forte, sont les cotylédons, l'albumen ou l'endo-sperme, et l'embryon.

Le fragment promoteur est de préférence utilisé pour diriger une expression constitutive, soit dans le cytoplasme, soit dans les mitochondries, soit dans les deux à la fois. Pour obtenir l'expression d'une protéine hétérologue dans les mitochondries, la séquence codant pour le peptide signal mitochondrial de l'aconitase est fusionnée à l'extrémité 5' de la séquence codante.

Il peut également être envisagé d'utiliser des fragments promoteurs s'étendant plus en amont du site HindIII (-1293). Les inventeurs ont identifié un clone par la technique RACE dont l'extrémité 5' sur la séquence génomique se situe en position -52 en amont de la première TATA box. Ceci laisse suggérer qu'une initiation de la transcription peut avoir lieu plus en amont. Ces résultats indiquent l'existence éventuelle d'un deuxième promoteur qui pourrait être responsable de la transcription de l'une des deux formes de

l'aconitase à partir du gène. Ce deuxième promoteur fait également partie de l'invention.

Le promoteur du gène de l'aconitase de maïs fait également partie de l'invention. Il est compris dans la partie amont de la séquence génomique de la figure 18, notamment au sein du fragment SalI-HindIII cité dans l'exemple 12.

Les régions promotrices de l'invention sont normalement utilisées sous forme de gène chimérique. Ces gènes chimériques comportent :

- i) une région promotrice selon l'invention, et
- ii) une séquence hétérologue transcrite, codante ou non, placée sous le contrôle dudit promoteur.

L'expression "séquence hétérologue transcrite, codante ou non" signifie dans le contexte de la présente invention toute séquence transcrite autre que celles normalement associées avec le promoteur de l'aconitase dans la plante. Cette expression englobe des séquences non-codantes telles que des ribozymes ou des séquences antisens, dirigées contre un transcrit d'intérêt, ainsi que des séquences codantes.

Comme exemples de séquences codantes, on peut citer :

- des gènes rapporteurs tel que Gus codant pour le Beta glucuronidase ;
- des séquences conférant une résistance à un antibiotique, par exemple la résistance à la kanamycine, la gentamycine, le G418, etc... ;
- toutes séquences codantes conférant des traits intéressants du point de vue agronomique, par exemple des séquences impliquées dans la stérilité mâle, par exemple des gènes mitochondriaux non-édités, des séquences conférant une résistance à un virus ou aux herbicides etc. ;

- des séquences codantes de gène de mammifères, par exemple l'interféron, l'hormone de croissance, l'insuline et tout autre produit susceptible d'être utile en tant que médicament.

Ces gènes chimériques permettent d'obtenir une forte expression dans les graines sèches.

L'invention concerne également des gènes chimériques capable d'être exprimés chez les plantes caractérisés en ce qu'ils comportent :

i) un promoteur fonctionnel chez les plantes autre que celui naturellement associé avec le gène de l'aconitase, et

ii) une séquence codant pour la protéine de l'invention, telle que définie ci-dessus, ou pour un fragment protéique, tel que défini ci-dessus.

Comme exemples de promoteur dans ce type de gène chimérique, on peut citer le promoteur du 35S, le promoteur NOS, des promoteurs semences spécifiques de la plante.

Comme promoteurs spécifiques de graines, on peut citer le promoteur du gène de la napin (EP-A-0255378), ainsi que les promoteurs des gènes AT2S d'Arabidopsis thaliana, c'est-à-dire les promoteurs PAT2S1, PAT2S2, PAT2S3 et PAT2S4 (Krebbers et al., 1988, Plant Physiol., 87 : 859-866).

La séquence codante correspond à celle codant pour la protéine de l'invention présentant une activité d'aconitase ou pour un fragment protéique présentant soit une homologie immunologique avec l'aconitase ou une activité d'aconitase.

L'invention concerne en outre des gènes chimériques capable d'être exprimés chez les plantes caractérisés en ce qu'ils comportent :

i) un promoteur fonctionnel chez les plantes,

ii) une séquence transcrite, codante ou non, placée sous le contrôle dudit promoteur,

iii) un ou plusieurs introns du gène de l'aconitase végétale, tel(s) que défini(s) ci-dessus.

La présence d'un ou plusieurs introns stabilise l'ARNm transcrit. Le ou les intron(s) peuvent être positionné(s) au sein de la séquence transcrite, ou peuvent être placé(s) entre le promoteur et le début de la séquence transcrite, comme indiqué ci-dessus.

L'invention vise également des plasmides ou vecteurs caractérisés en ce qu'ils contiennent au moins l'une des séquences d'acide nucléique selon l'invention. De préférence, les plasmides et vecteurs permettent la transformation stable de la plante, par exemple les plasmides Ti d'Agrobacterium, des vecteurs viraux tels que les Geminivirus ou le CaMV.

L'invention concerne également des cellules végétales et des plantes transgéniques transformées de manière stable par une séquence d'acide nucléique selon l'invention. Comme exemple de plantes transgéniques, on peut citer les espèces appartenant aux familles botaniques telles que les légumineuses (par exemple les haricots, pois, etc...), les crucifères (par exemple, les choux, radis, etc...), les solanacées (par exemple les tomates, pommes de terre, etc...) les cucurbitacées (par exemple le melon), les chénopodiacées (par exemple la betterave potagère), et les ombellifères (par exemple, les carottes, céleris, etc...). Les cucurbitacées et les Brassicacées sont particulièrement préférés. On peut également citer les céréales telles que le blé, le maïs, l'orge, le triticale et le riz, et les oléagineux tels que le tournesol, le soja et le colza.

Tous les moyens connus pour introduire de l'ADN étranger dans des plantes peuvent être utilisés, par

exemple Agrobacterium, électroporation, fusion de protoplastes, bombardement avec canon à particules, ou pénétration d'ADN dans des cellules comme le pollen, la microspore, la graine et l'embryon immature, vecteurs viraux tels que les Geminivirus ou les virus satellites. Agrobacterium tumefaciens et rhizogenes constituent le moyen préféré. Dans ce cas, la séquence de l'invention est introduite dans un vecteur approprié avec toutes les séquences régulatrices nécessaires telles que promoteurs, terminateurs, etc... ainsi que toute séquence nécessaire pour sélectionner les transformants.

L'invention vise également des anticorps monoclonaux ou polyclonaux capable de reconnaître, de préférence de manière spécifique, la protéine et les fragments protéiques de l'invention. Ces anticorps sont produits selon les techniques habituelles dans l'art en utilisant, comme immunogène, la protéine purifiée selon les techniques décrites dans les exemples ci-dessous ou la protéine recombinante. La technique de purification de l'exemple 1 ci-dessous conduit à la production de deux formes de l'aconitase (appelées AcoI et AcoII). En utilisant l'une de ces deux protéines comme immunogène, il est possible, selon l'invention, de produire des anticorps monoclonaux capable de faire la distinction entre les deux formes. Ces anticorps monoclonaux spécifiques constituent des sondes immunologiques utiles pour étudier la localisation subcellulaire des deux formes enzymatiques, ou pour la purification.

De manière générale, les anticorps de l'invention peuvent être utilisés pour le crible de banques d'ADNC ou pour d'autres utilisations telles que l'étude de la parenté entre les aconitases de différentes espèces.

Les protéines et acides nucléiques de l'invention trouvent de nombreuses applications. Selon un mode de réalisation particulièrement préféré, les séquences d'acide nucléique et les gènes chimériques de l'invention sont utilisés dans un procédé pour la modification du métabolisme d'une plante. Ce procédé est caractérisé par l'introduction, dans la plante, d'un insert génétique comportant une séquence d'acide nucléique de l'invention.

La modification du métabolisme de la plante peut consister, par exemple :

- en la modification de la quantité d'aconitase produite par la plante ;
- en une modification de l'expression temporelle ou spatiale de l'aconitase ;
- en une modification de la nature de l'aconitase produite.

De préférence, la modification consiste en une modification de la quantité d'aconitase produite par la plante. Il peut s'agir d'une augmentation ou d'une inhibition de la production d'aconitase.

Une augmentation est obtenue par l'introduction dans la plante d'un insert génétique comprenant un promoteur fonctionnel chez les plantes, et sous le contrôle de ce promoteur, une séquence codant pour la protéine de l'invention. Le promoteur peut être celui naturellement associé avec le gène de l'aconitase végétale ou un promoteur hétérologue au gène de l'aconitase. Lorsque le promoteur est le promoteur naturel, l'insertion de la séquence codante avec son propre promoteur conduit à une surproduction d'aconitase, l'expression des copies supplémentaires du gène de l'aconitase étant spatialement et temporellement identique au gène endogène. En revanche, si le promoteur est hétérologue au gène de

l'aconitase, il peut s'agir d'un promoteur constitutif ou encore d'un promoteur spécifique de certains tissus ou de certains stades de développement. Ceci permettra d'obtenir une expression temporelle et spatiale différente de celle produite par le gène endogène.

La séquence codante utilisée pour obtenir une surproduction d'aconitase peut être la séquence codante entière exemplifiée par celle illustrée dans la figure 1, ou peut correspondre à l'une des deux formes (mitochondriale ou cytosolique) permettant d'obtenir la surproduction dans une localisation subcellulaire définie. Les séquences codant pour les protéines présentant au moins 75 % d'homologie avec celle de la figure 1 peuvent également être utilisées.

Selon cet aspect de l'invention, la surproduction d'aconitase conduit à une surproduction d'au moins l'un des acides organiques du cycle de Krebs ou du cycle glyoxylate, notamment le citrate, l'isocitrate, le succinate, le glyoxylate, le malate, l'oxaloacétate, l' α -cetoglutarate, le fumarate.

Cette variante de l'invention constitue donc un procédé pour obtenir une surproduction d'acides organiques chez les plantes. Les acides ainsi obtenus peuvent ensuite être extraits de la plante, en faisant appel à des méthodes connues en soi, pour l'utilisation dans l'industrie agroalimentaire, par exemple en tant qu'additif, antioxydants etc. La production de citrate est particulièrement préférée, par exemple en tant qu'additif pour des jus de fruits.

La quantité d'aconitase produite par la plante peut également être diminuée par l'utilisation d'un ribozyme ou d'une séquence antisens dirigé contre au moins une partie du transcrit du gène de l'aconitase. Les ribozymes et les séquences antisens ont la structure décrite plus haut. Ces séquences

inhibitrices peuvent être utilisées dans un gène chimérique selon l'invention, le promoteur étant choisi afin de permettre une expression appropriée. Par exemple, un promoteur inductible permettrait la production du ribozyme et par conséquent l'inactivation du transcrit de l'aconitase uniquement lorsque les conditions d'induction étaient remplies. Le ribozyme ou la séquence antisens peut également être sous le contrôle du promoteur de l'invention.

Selon cette variante de l'invention, l'inhibition de la production d'aconitase conduit à une diminution de la production des acides organiques du cycle de Krebs et du cycle glyoxylate. Par conséquent, il y a augmentation de la quantité d'Acétyl coenzyme-A et d'acide citrique disponibles. L'Acétyl coenzyme-A étant le produit du catabolisme lipidique, protéique et polysaccharidique, il s'ensuit une modification du catabolisme de ces substances ou de certaines de ces substances. Cet aspect de l'invention peut donc être appliqué par exemple à la réorientation du métabolisme ou du catabolisme d'amidon, d'huiles, de protéines, etc, pour les industries agro-alimentaires, chimiques, pharmaceutiques et cosmétologiques.

Selon une autre variante de l'invention, la distribution tissulaire de l'aconitase peut être modifiée. Par exemple, l'insertion de copies supplémentaires de l'une des deux formes d'aconitase, par exemple la forme mitochondriale plutôt que l'autre forme, conduit à la surproduction de cette forme.

En outre, le clivage systématique du peptide signal mitochondrial, par l'utilisation d'un ribozyme telle que définie ci-dessus, conduit à la production de la forme cytosolique seulement de l'enzyme.

Bien entendu, l'utilisation d'un promoteur spécifique de certains tissus permettra l'obtention

d'aconitase dans les tissus où l'expression n'est pas normalement très forte.

Il peut également être envisagé de modifier la nature de l'aconitase produite par une plante par l'insertion, dans une plante appartenant à une première espèce, d'une séquence codant pour l'aconitase d'une deuxième espèce. Ceci peut être intéressant lorsque la deuxième espèce présente des caractéristiques particulièrement avantageuses, par exemple lorsque l'aconitase d'une espèce particulière est moins sensible à certaines substances, ou lorsque son expression est réglée différemment de celle de la plante hôte.

La modification du métabolisme de la plante par une augmentation ou une diminution de la production d'aconitase peut également influencer sur le métabolisme des ions métalliques, tels que le fer, le cadmium et le cuivre. Ces effets sont vraisemblablement exercés grâce au rôle d'IRE-BP que peut jouer l'aconitase végétale.

Différents aspects de l'invention sont illustrés dans les figures :

- Figure 1 : séquence nucléotidique du gène codant pour une protéine à activité aconitase chez Arabidopsis thaliana.

Légende : lettres minuscules = séquence non codante, lettres majuscules = séquence codante, acides aminés en caractères gras = séquence montrant une homologie avec les séquences protéiques NH₂ terminales de l'aconitase mitochondriale de pomme de terre et des aconitases (AcoI et AcoII) de graines de melon.

- Figure 2 : comparaison des séquences protéiques des aconitases de différentes espèces. Les séquences protéiques des IRE-BP et des aconitases de E. coli, de porc et des levures ont été extraites de la banque

"Swiss Prot 26" (version Aout 93). L'alignement a été réalisé avec le programme PC Gene Clustal (Intelli Genetics).

Légende : ARABACO = Aconitase d'Arabidopsis thaliana, MELONACO = Aconitase de melon (Cucumis melo), IREB = Iron Responsive Element Binding Protein, ACON = Aconitase, ∇ = cystéines constituant le centre Fer-Soufre, ↓ = acides aminés formant le site actif, * = acides aminés identiques, . = acides aminés similaires.

- Figure 3 : séquence nucléotidique et protéique du clone d'ADNc n° 1 d'Arabidopsis thaliana codant pour une protéine à activité aconitase.

Légende : lettres minuscules = séquence non codante, lettres majuscules = séquence codante, acides aminés en caractères gras = séquence montrant une homologie avec les séquences protéiques NH₂ terminales de l'aconitase mitochondriale de pomme de terre et des aconitases (AcoI et AcoII) des graines de melon.

- Figure 4 : séquence nucléotidique et protéique du clone d'ADNc n° 16 de melon codant pour une protéine à activité aconitase.

Légende : lettres minuscules = séquence non codante, lettres majuscules = séquence codante

- Figure 5 : cycle de Krebs

- Figure 6 : cycle du Glyoxylate

- Figure 7 : cycle de Krebs = source de squelettes carbonés utilisés dans plusieurs voies de biosynthèse.

- Figure 8 : compartimentation des enzymes impliquées dans le cycle du Glyoxylate.

1 : citrate synthase, 2 : isocitrate lyase, 3 : malate synthase, 4 : malate déshydrogénase.

- Figure 9 : schéma du principe de la RACE-PCR.

Légende : am et am' = amorces complémentaires.

- Figure 10 : carte de restriction déduite du clone ADNc n° 1 d'Arabidopsis thaliana codant pour une protéine à activité aconitase.

B : bamHI, E : EcoRI, H : HindIII, RV : Eco RV, Pv : PvuII.

- Figure 11 A : séquence nucléotidique et protéique de l'extrémité 5' du cDNA n° 1 d'Arabidopsis thaliana codant pour une aconitase.

Caractères gras : séquence protéique montrant de l'homologie avec les extrémités NH₂ terminales des aconitases de pomme de terre et de melon.

- Figure 11 B : comparaison des séquences NH₂ terminales de l'aconitase déduite du cDNA d'Arabidopsis thaliana, de l'aconitase de melon (Cucumis melo), et de l'aconitase mitochondriale de pomme de terre (Solanum tuberosum).

- Figure 12 : pourcentages d'homologies entre les aconitases de différentes espèces.

Légende : ARABACO = Aconitase d'Arabidopsis thaliana, MELONACO = Aconitase de melon (Cucumis melo), IREB = Iron Responsive Element Binding Protein, ACON = Aconitase, le chiffre gras = pourcentage de résidus identiques, le chiffre entre parenthèses = pourcentage de résidus similaires.

- Figure 13 : arbre phylogénique des aconitases de différentes espèces.

Légende : ARABACO = Aconitase d'Arabidopsis thaliana, MELONACO = Aconitase de melon (Cucumis melo), IREB = Iron Responsive Element Binding Protein, ACON = Aconitase.

- Figure 14 : cartographie physique du gène de l'aconitase d'Arabidopsis thaliana.

A : carte de restriction de la partie 5' du fragment génomique isolé à partir de la banque DASH,

B : carte de restriction de la partie 3' du fragment génomique isolé à partir de la banque EMBL,

C : cartographie du gène de l'aconitase déduite des deux cartes précédentes.

B = BamHI, Ps = PstI, Pv = PvuII, H = HindIII, E = EcoRI, X = XbaI, DASH = vecteur phagique DASH, EMBL = vecteur phagique EMBL3.

- Figure 15 : structure IRE-like putative (position -85 à -52 sur la séquence génomique).

- Figure 16 : détail de la fusion traductionnelle du promoteur de l'aconitase avec le gène rapporteur de la β -glucuronidase portée par le vecteur binaire pBIOS 170.

a) Séquence de la région en amont du codon d'initiation putatif (Met) de l'aconitase.

b) Séquence de la région en amont du codon d'initiation de la β -glucuronidase (vecteur pBI 101.1).

c) Séquence de la fusion du promoteur de l'aconitase avec le gène de la β -glucuronidase (tataa = Tata box).

Caractères standards : séquence de l'aconitase, caractères en italique : séquence de pBI 101.1.

- Figure 17 : alignement des séquences protéiques et nucléotidiques d'A. thaliana (acides aminés 240 à 269) avec la partie correspondante du maïs. La séquence de maïs a été obtenue par P.C.R., utilisant comme amorces, oligo D III et oligo D IV. Une homologie nucléique de 77.5%, et protéique de 87.5% était détectée. La séquence d'une centaine de bases a ensuite servi de sonde sur de l'ADN génomique de maïs et a permis l'identification du gène.

- Figure 18 : 6036 paires de bases de la séquence génomique d'un gène aconitase de maïs obtenues par séquençage du clone R1. La séquence protéique à partir

d'exon 2 est également indiquée, bien que l'extrémité 5' de l'exon 2 n'a pu être déterminé avec certitude. L'exon 1 et l'intron 1 n'ont pas été identifiés, mais ils se situent dans les quatre mille premières bases d'un fragment d'une taille de 6800 paires de bases SalI-HindIII du clone génomique R1 ; HindIII étant à la position 2841 de la séquence présentée.

- Figure 19 : alignement des aconitases Arabidopsis et maïs attestant la très grande conservation de ces enzymes. (* = acides aminés identiques, . = acides aminés similaires).

- Figure 20 : alignement entre 3 aconitases végétales identifiées dans le cadre de la présente invention (ZMACO = Aconitase maïs, ATACO = aconitase Arabidopsis, CMACO = aconitase melon, * = acides aminés identiques, . = acides aminés similaires).

- Figure 21 : amorces D I à D VI utilisées dans l'exemple 11.

EXEMPLES

EXEMPLE I : PURIFICATION DE 2 POLYPEPTIDES AYANT UNE ACTIVITE ACONITASE A PARTIR DE GRAINES DE MELON :

L'aconitase a été purifiée à partir de graines de melon par des techniques chromatographiques traditionnelles.

L'extrait brut est préparé par broyage de 100 g de graines au waring blendor (3 mn, 4°C, vitesse maximale), puis agité dans 100 ml de tampon Imidazole 20 mM ; pH 7,5 (1 h, 4°C). La fraction soluble contenant l'activité est récupérée après centrifugation (13 000 g, 30 mn, 4°C), puis filtrée

sur une feuille de miracloth permettant d'éliminer les débris cellulaires et une grande partie des lipides.

La première étape chromatographique consiste en une chromatographie d'échange d'ions sur une résine Q sépharose fast flow (PHARMACIA). La suspension protéique brute (1 360 mg) est déposée au sommet de la colonne équilibrée avec un tampon d'Imidazole 20 mM, pH 7,5. Les protéines sont éluées en conditions isocratiques dans un tampon Imidazole 20 mM, pH 7,5 additionné d'acétate de sodium 200 mM.

Les fractions actives (260 mg de protéines) sont dessalées sur des colonnes EconoPac 10 DG (BIORAD) avec un tampon Imidazole 20 mM, pH 7,5. Les protéines sont alors déposées sur une colonne contenant une résine d'affinité sur colorant Yellow 86 (SIGMA). L'élution est ensuite réalisée par un gradient linéaire d'acétate de sodium allant de 0 à 0,5 M dans un tampon Imidazole 20 mM, pH 7,5.

Les fractions actives sont de nouveau récupérées (14,4 mg de protéines) puis ajustées à une concentration en sulfate d'ammonium de 1,2 M.

Les extraits protéiques sont alors déposés sur une colonne d'interactions hydrophobes Protein Pack HIC Phényl 5 PW (WATERS). L'élution des protéines est ensuite effectuée par un gradient linéaire de sulfate d'ammonium allant de 1,2 à 0 M dans un tampon Imidazole 20 mM, pH 7,5.

Deux pics d'activité sont alors détectés. Le premier pic d'activité (Aco I) est élué pour une concentration en sulfate d'ammonium voisin de 0,9 M. Le deuxième pic d'activité (Aco II) est élué pour une concentration plus faible de 0,7 M.

Les fractions actives sont de nouveau dessalées (fraction Aco I : 0,76 mg ; fraction Aco II : 0,44 mg) dans un tampon Imidazole 20 mM, pH 7,5. Enfin, les

fractions actives de chacun des pics d'activité sont déposées sur une colonne d'échange d'ions Mono Q MR 5/5 (PHARMACIA).

Cette dernière chromatographie permet d'obtenir deux protéines à activité aconitase relativement pures après une élution avec un gradient linéaire d'acétate de sodium de 0 à 0.3 M dans un tampon d'Imidazole 20 mM, pH 7,5. Les protéines Aco I et Aco II étaient éluées pour une concentration en acétate de sodium voisine de 200 mM. Les protéines Aco I et Aco II en fin de purification ont été purifiées d'un facteur respectivement de 878 fois et 763 fois. Ces deux protéines montrent des activités spécifiques voisines de 1 518,42 mmol/mn/mg pour Aco I et de 1 321,15 mmol/mn/mg pour Aco II.

90 mg de protéine pure Aco I et 60 μ g d'Aco I sont obtenus en fin de purification.

Tableau I :

Etapas de purification	Protéines (mg)	Activité totale (nmol/min)	Activité spécifique (nmol/min/mg)	Activité restante (%)	Facteur de Purification
Extrait brut	1360	2360	1.73	100	-
Sépharose QFF	266	1565	5.88	66.3	3.4
Yellow 86	14.4	725	50.34	30.7	29.2
Phényl HIC					
- fraction Aco I	0.76	229	781.78	9.7	451.9
- fraction Aco II	0.44	54	122.48	2.3	70.8
Mono Q					
- fraction Aco I	0.09	137	1518.42	5.8	877.7
- fraction Aco II	0.06	79	1321.15	3.3	763.7

**EXEMPLE 2 : DETERMINATION DE LA SEQUENCE NH2 TERMINALE
DES 2 ACONITASES :**

Un gel de polyacrylamide (gel de séparation 10 % ; gel de concentration 5 %) est coulé 24 h avant son utilisation (système Mini Protean II BIORAD). Une pré-électrophorèse de 1 h 30 à 20 mA permet d'homogénéiser le gel.

Les échantillons ne doivent pas être dénaturés, ils sont placés 15 mn à 37°C. La séparation se fait pendant 45 mn à 190 V.

Après la séparation, les protéines sont transférées sur une membrane PVDF (Poly Vinyl DiFluoride). La membrane est plongée 10 s dans une solution de méthanol 100 % puis incubée dans le tampon de transfert CAPS (3[cyclohexylamino]t-propane sulfonic acid) 10 mM, pH 11, additionné de méthanol 10 % jusqu'à l'équilibre.

L'électrotransfert des protéines se fait en condition semi-sèche avec le système multiphor 2117 LKB pendant 1 h à 60 mA (0,8 mA/cm² de membrane).

Les protéines transférées sur la membrane PVDF sont visualisées après coloration au bleu de Coomassie R250 (solution à 0,1 % dans du méthanol 50 %) pendant 5 mn, puis décoloration 5 à 10 mn dans la solution suivante méthanol, acide acétique, eau (5 : 10 : 40 vol).

La partie de membrane correspondant à la protéine intéressante est découpée, et la protéine séquencée.

La partie NH₂ terminale de la protéine est séquencée selon la réaction de HEWICK et al (1981) avec le séquenceur automatique Applied-Biosystem.

La séquence NH₂ terminale des aconitases de graines de melon, que les inventeurs ont déterminée, a

pu être comparée (Figure 11) avec la séquence NH₂ terminale de l'aconitase mitochondriale de pomme de terre (COURTOIS-VERNIQUET et DOUCE, communication personnelle). Une forte homologie entre ces séquences est constatée : sur 14 résidus séquencés, 9 sont communs avec l'aconitase mitochondriale de pomme de terre. Cela peut signifier que l'aconitase mitochondriale de melon a été purifiée ou bien alors que les séquences NH₂ terminales des aconitases mitochondriales et cytosoliques sont très homologues. L'absence de méthionine comme premier acide aminé pourrait indiquer une maturation de la protéine lors de l'importation dans les mitochondries (clivage du peptide signal).

EXEMPLE 3 : FABRICATION D'ANTICORPS POLYCLONAUX DIRIGES CONTRE LES ACONITASES :

Les lapins (hybrides femelles de 2 à 3 mois) sont immunisés par voie intradermique avec 10 µg d'aconitase purifiée, émulsionnée à un volume égal d'adjuvant de Freund complet lors de la première immunisation, incomplet pour les immunisations qui suivront (40 et 70 jours après la première injection).

Les prélèvements de sang, pratiqués dans la veine centrale de l'oreille sont effectués régulièrement (toutes les semaines). Le sang recueilli sur héparine est rapidement centrifugé (3000 g, 4°C, 5 mn). Les plasmas contenant les anticorps sont répartis en aliquots et conservés à -20°C.

La caractérisation des anticorps polyclonaux se fait par la technique de Western Blot selon le protocole détaillé ci-dessous.

L'électrophorèse des protéines ou des fractions purifiées est effectuée en gel dénaturant SDS-PAGE 10 % (gel de séparation), gel de concentration 3 %.

Les échantillons sont préparés dans le tampon suivant : Tris/HCl 10 mM ; pH 8,0 ; EDTA 1 mM ; SDS 2,5 % ; β -mercaptoéthanol 5,0 % ; bleu de bromophénol 0,01 %. Les protéines sont dénaturées par chauffage (100°C, 5 mn).

Un dépôt de 100 μ l d'extrait brut (préparation : voir exemple I) traité dans les conditions décrites ci-dessus en présence de 10 % de sucrose, est réalisé dans un puits unique.

Les marqueurs protéiques colorés de poids moléculaire (Rainbow-AMERSHAM) sont les suivants : myosine (200 000) phosphorylase b (97 400), inhibiteur de trypsine (21 500), lysozymes (14 300). L'électrophorèse s'effectue sous un courant de 20 mA par gel pendant 1 h.

Les protéines séparées sur gel SDS-PAGE sont transférées sur membrane de nitrocellulose. L'électrotransfert se fait avec l'appareil mini trans blot BIORAD ou le Phast System PHARMACIA (transfert semi-sec selon les recommandations du fabricant), en fonction du type de gel qui a permis la séparation préalable des protéines.

La révélation de la membrane par les anticorps se fait comme suit :

- la membrane est incubée dans une solution saturante TBST (Tris-HCl 20 mM ; NaCl 500 mM ; Tween 20 0,05 %) additionnée de Régilait (5 %) pendant un minimum de 3 h.

- puis la membrane est placée 2 h dans une solution d'incubation TBST contenant le premier anticorps (anticorps polyclonal dilué au 1/8000 ou anticorps monoclonal : surnageant de culture).

- l'élimination de l'excès d'anticorps se fait par trois lavages successifs de 3 mn dans une solution TBST. La membrane est ensuite incubée 2 h dans une solution TBST, en présence d'un second anticorps conjugué à la phosphatase alcaline (anti-immunoglobulines de lapin diluées au 1/9 000 ou anti-immunoglobulines de souris diluées au 1/4 000, SIGMA). Après cette étape, 3 nouveaux lavages sont effectués.

- enfin, le complexe antigène-anticorps est révélé dans la solution de développement suivante : 5 ml de tampon phosphatase alcaline (Tris/HCl 100 mM ; pH 9,5 ; NaCl 100 mM ; MgCl₂ 5 mM), 33 µl de nitro blue tetrazolium (NBT) à 50 mg/ml préparé dans du diméthylformamide 70 %, 16,5 µl de 5-Bromo-4-chloro-3-indolylphosphate (BCIP) à 50 mg/ml préparé dans de la diméthylformamide.

Des anticorps polyclonaux ont été produits après immunisation de deux lapins avec la protéine Aco I (10 µg/injection) et de deux autres lapins avec la protéine Aco II (10 µg/injection).

La réactivité et la spécificité des anticorps sont testées sur des extraits bruts ou purifiés de protéines de graines de melon séparées par électrophorèse et transférées sur une membrane de nitrocellulose. Les anticorps donnent une réponse forte pour une dilution jusqu'au 1/8 000. Les deux plasmas donnent le même type de réponse.

Dans l'extrait brut, une seule bande est reconnue à environ 97 kDa, démontrant la spécificité des anticorps polyclonaux. De même, dans les extraits protéiques des autres étapes chromatographiques, une seule bande est reconnue. Notons que les plasmas reconnaissent aussi bien Aco I que Aco II et ce, quel que soit le matériel ayant servi à l'immunisation (Aco I ou Aco II).

La forte spécificité de ces anticorps permettra leur utilisation principalement pour le crible de banques ADNc ou pour d'autres utilisations éventuelles (localisation cytotogique de l'aconitase, parenté entre les aconitases de différentes espèces).

EXEMPLE 4 : ISOLEMENT D'ADNc ACONITASE PAR IMMUNOSCREENING D'UNE BANQUE D'EXPRESSION DE FRUIT MATURE DE MELON :

Les inventeurs disposaient d'une banque ADNc de fruit de melon mûr (Cantaloup charentais).

Les ADNc sont clonés au début du gène de la β -galactosidase dans les phages UNI-ZAP. Un clonage orienté permet de réduire de moitié les clones qui ne sont pas dans la bonne phase de lecture.

Le répresseur lacI q qui bloque la transcription du gène lac Z est apporté par la bactérie XL1 Blue. Ce répresseur est important car il permet de contrôler l'expression des protéines de fusion qui peuvent être toxiques pour E.coli.

Ainsi, la production de protéines de fusion se fait au moment choisi grâce à l'addition de l'IPTG (isopropylthio β -D galactoside) qui est un inducteur du promoteur du gène lac Z.

Les étalements de la banque se font comme décrit dans Sambrook et al (1989, Molecular Cloning : a Laboratory Manual, Cold Spring Harbor Laboratory Press, New-York). Au bout d'environ 8 h, lorsque les plages de lyse apparaissent, les filtres préalablement imbibés dans une solution d'IPTG 10 mM puis séchés, sont déposés à la surface des boîtes. Les boîtes sont de nouveau incubées pendant 3 h à 37°C permettant à l'IPTG d'induire la protéine de fusion qui se fixera sur le filtre après que les bactéries aient été

lysées. Ce sont les bactéries infectées formant le contour de la plage de lyse qui produiront la protéine de fusion. Ainsi, lors de la révélation par les anticorps, les plages de lyse apparaîtront sous forme de halo. Les filtres sont percés en trois positions asymétriques pour pouvoir repérer facilement les plages positives.

La révélation de la protéine de fusion se fait comme décrit dans l'exemple 3 à l'aide des anticorps polyclonaux produits.

Les clones positifs sont repérés en superposant les filtres révélés avec la boîte correspondante. Le morceau de gélose contenant la zone positive est prélevé à l'aide d'une pipette pasteur, puis placé dans 1 ml de tampon SM contenant une goutte de chloroforme (inhibant le développement bactérien) pendant 1 h à température ambiante pour permettre la diffusion des phages dans la solution. Les tubes sont conservés à 4°C.

Après le premier crible, les plages de lyse positives qui sont prélevées ne sont pas pures (plages confluentes). Il convient, pour isoler avec certitude un clone positif, d'effectuer plusieurs cribles successifs en réétalant les plages de lyse récupérées de façon à obtenir des plages de lyse isolées.

Sur les 4.10^5 phages qui ont été étalés, 28 clones purs reconnus par les anticorps polyclonaux ont alors pu être isolés.

Ces clones ont été analysés par digestion enzymatique, hybridation et séquençage.

La digestion de ces clones par EcoRI et Xho I montre deux catégories de clones :

- Les clones dont l'insertion contient un site de restriction EcoRI mais pas de site de restriction Xho I (21 clones).

- Les clones dont l'insertion contient un site de restriction Xho I mais pas de site de restriction EcoRI (7 clones).

Pour étudier la parenté de ces clones les inventeurs ont réalisé des hybridations en dot-blot (ceci permet d'effectuer un tri préliminaire sans avoir à effectuer de digestion et de transfert sur membrane).

Le fragment obtenu, après digestion par l'enzyme de restriction EcoRI, du clone le plus grand (2,6 Kb) et du clone le plus court (2,3 Kb) de la première catégorie ont servi de sonde. Les 21 clones de la première catégorie sont apparentés : ils donnent un signal lors de l'hybridation.

Après digestion par EcoRI et transfert sur membrane, les clones sont hybridés avec le fragment EcoRI de 2,2 Kb du 10 donnent un signal, les inventeurs ont pu déterminé 3 classes de clones. La première classe (I) est constituée de 11 clones dont l'insert EcoRI a une taille de 1,7 Kb. La deuxième classe (II) de clones (9 clones) comporte un insert EcoRI de 1,8 Kb. La 3ème classe (III) est constituée du clone le plus grand avec un insert EcoRI de 2,2 Kb (n°16).

EXEMPLE 5 : SEQUENCAGE DES ADNC DE MELON - HOMOLOGIES AVEC LES IRE-BP :

La séquence complète du clone le plus long (n° 16) a été déterminée par la technique classique de Sanger et al. Elle est présentée dans la figure 4.

En 3', la séquence fait apparaître une queue poly A de 80 résidus. Un clone de chacune des classes I et II a été séquencé en 3' et en 5' montrant une homologie parfaite avec le clone 16 mais avec des

queues poly A d'une longueur moins importante (respectivement 21 et 27 résidus). D'après toutes nos données d'analyses de restriction (simple et double digestion), d'hybridation et de séquençage, tous ces clones semblent provenir d'un même messager.

La séquence 5' du clone le plus long (n°16) n'a pas permis de retrouver la séquence protéique correspondant à l'extrémité NH₂ terminale. Ceci n'est pas étonnant car une protéine d'environ 100 kDa devrait être codée par un ARNm d'au moins 3 Kb. Ils se sont alors servis de cette séquence de 900 pb pour interroger les banques de données (SWISSPROT - CITI 2) et s'assurer qu'ils avaient bien isolé le clone ADNc correspondant à la protéine purifiée. Des homologies importantes ont été retrouvées avec l'ensemble des séquences d'aconitase ou d'IRE-BP, en particulier avec l'IRE-BP humaine et l'aconitase d'E. coli (environ 60 %).

Ces résultats suggéraient que l'ADNc correspondant à la protéine purifiée avait bien été isolé. Cependant, l'extrémité 5' de cet ADNc n'était pas présente.

Avec des fragments 5' terminaux, les inventeurs ont essayé d'isoler des clones complets dans cette banque, mais sans succès. Aussi ont-ils pris la décision de cribler une nouvelle banque d'ADNc qui provenait d'une autre espèce végétale.

EXEMPLE 6 : ISOLEMENT DE L'ADNc DE L'ACONITASE D'ARABIDOPSIS THALIANA ET COMPARAISON DES SEQUENCES PROTEIQUES DES ACONITASES DE DIFFERENTES ESPECES (ANIMALES, VEGETALES, BACTERIENNES) :

Comme Arabidopsis thaliana est aujourd'hui l'espèce modèle par excellence (petit génome, cycle de vie court, cartes génétiques et RFLP disponibles, etc...), il a été choisi de cribler une banque ADNc de siliques immatures (écotype Columbia), construite par GIRAUDAT et al (1992, Isolation of the Arabidopsis ABI3 gene by Positional cloning, The Plant Cell 4 : 1251-1261). Cette banque a été réalisée avec le vecteur Lambda ZAP II de STRATAGENE. Le clonage des ADNc n'est pas orienté (clonage EcoRI).

La difficulté dans ce choix de stratégie est de trouver les bonnes conditions d'hybridation car les inventeurs disposent d'un clone ADNc servant de sonde d'une espèce donnée (melon) pour cribler une banque d'une autre espèce (Arabidopsis thaliana). Ils ne connaissent pas le degré de conservation de l'aconitase chez les plantes. Aussi, plusieurs conditions d'hybridation dans lesquelles la concentration en formamide variait (0, 10, 20, 30, 40, 50 %) ont été testées. De même, des conditions de lavages progressifs ont été utilisées.

Pour chaque condition d'hybridation, environ $1,5 \cdot 10^5$ phages ont été étalés. La sonde utilisée est le fragment d'ADN de 2,2 Kb obtenu après digestion par l'enzyme de restriction EcoRI du clone ADNc de melon n°16.

7 clones hybridant avec la sonde dans les conditions suivantes ont été isolés : hybridation effectuée en présence de 30 % de formamide à 42°C et 2 lavages dans une solution SSC 2X pendant 30 mn à 42°C.

Ces clones ont été analysés après digestion par l'enzyme de restriction EcoRI. La carte de restriction sommaire des clones les plus longs, 3 clones de 3,2 Kb, est décrite dans la figure 10.

Les 4 autres clones sont identiques, leur taille est de 2,3 Kb. La carte de restriction montre qu'ils appartiennent à la même classe que les clones les plus longs, seul le fragment EcoRI de 0,8 Kb à l'extrémité 5' est absent (ces clones proviennent vraisemblablement d'une méthylation incomplète des sites EcoRI internes des ADNc).

Les 3 clones de 3,2 Kb (n° 1, 2, 3) ont été séquencés à chaque extrémité (amorces universelles T3 et M13). A l'extrémité 3', il a été observé une queue poly A constituée respectivement de 40, 15 et 16 résidus. La séquence de l'extrémité 5' (Figure 11) a montré une forte homologie avec la séquence NH2 terminale protéique de l'aconitase de melon (9 acides aminés identiques sur 14) et de l'aconitase mitochondriale de pomme de terre (8 acides aminés identiques sur 14) (figures 10 et 3).

Un clone ADNc correspondant à l'aconitase d'Arabidopsis thaliana avait donc vraisemblablement été isolé.

Les fragments EcoRI-EcoRV de 0,7 et 0,9 Kb (voir carte de restriction, Figure 10) ont été sous clonés dans le vecteur pBSII SK+ puis séquencés aux extrémités avec les amorces universelles T3 et M13. Le clone ADNc en totalité a par la suite été séquencé (Figure 3) par la stratégie de progression par oligonucléotides. Ce clone a une longueur totale de 3 210 pb. Le cadre de lecture ouvert a une longueur de 2769 pb. Le codon de terminaison est un codon TAA (codon ochre), la partie non traduite en 3' a une longueur de 441 pb. A l'extrémité 5', il n'y a pas de

codon stop (cadre de lecture ouvert) ce qui pourrait laisser supposer la présence d'un peptide signal. A l'exception de la méthionine présente dans la séquence NH₂ terminale protéique, il n'a pas été observé d'autre méthionine plus en amont, ce qui indiquerait que ce clone ne soit pas totalement complet. Le poids moléculaire de la protéine à activité aconitase codé par cet ADNc déduit à partir de la séquence nucléotidique est de 98 490 Da. Le pI calculé est de 6,0. Ces valeurs sont en accord avec les résultats obtenus lors de la purification (PM : 97 kDa ; pI : 5,2).

Les essais d'obtention d'un clone plus long ont été infructueux et la séquence totale du clone le plus long est présentée dans la figure 3 (ADNc n° 1).

La séquence des extrémités 5' et 3' des ADNc n° 2, 3, 4 et 5 permettent leur positionnement sur la séquence du clone ADNc n° 1.

L'extrémité 5' des clones n° 2 et 3 se situe en position 66. Pour les clones 4 et 5 l'extrémité 5' est en position 134.

A l'extrémité 3', les clones 2, 3, 4 et 5 se situent respectivement en position 2964, 2964, 2926, 2976. Les clones 2 et 3 sont strictement identiques ; même départ et même arrêt. L'analyse de la partie 3' non codante du clone ADNc n° 1 fait apparaître trois sites de polyadénylation possibles suivant le consensus AATAAA en position 2848, 2940 et 3026. Ces trois sites doivent être fonctionnels car 3 positions d'arrêt compatibles ont été obtenues pour les clones ADNc étudiés.

Les comparaisons de séquences de l'IRE-BP humaine et de l'aconitase mitochondriale de porc (ROUAULT et al, 1991 (supra) ; HENTZE et ARGOS, (1991, Homology between IRE-BP, a regulatory RNA-binding Protein,

Aconitase, and isopropylmalate isomerase, Nucleic Acids Research 19 : 1739-1740) avaient montré la parenté entre ces deux protéines (30 à 33 % d'identité).

L'isolement des ADNc chez Cucumis melo et Arabidopsis thaliana codant pour une aconitase a permis de rechercher les homologies existant avec d'autres aconitases de différentes espèces (Figure 2).

Les alignements des séquences protéiques montrent des régions fortement conservées. Les études cristallographiques de LAUBLE et al (Crystal Structures of Aconitase with Isocitrate and Nitroisocitrate bound, Biochemistry 31 : 2735-2748) en 1992 avaient permis d'identifier 23 résidus formant le site actif. Ces résidus sont conservés dans toutes les aconitases étudiées. Ces alignements font également apparaître que les protéines à activité aconitase étudiées chez Arabidopsis thaliana et Cucumis melo ont une homologie plus importante avec les IRE-BP.

La Figure 12 indiquant les pourcentages d'homologie entre les différentes aconitases montre une homologie très importante d'environ 90 % entre les aconitases végétales étudiées. Ces protéines apparaissent plus proches des IRE-BP (homologie d'environ 70 %) et de l'aconitase de E. coli (homologie d'environ 65 %) que des aconitases mitochondriales (homologie d'environ 45 %).

L'arbre de distance (Figure 13) permet de visualiser rapidement les parentés entre ces différentes aconitases.

Il a été montré qu'il existe vraisemblablement chez les plantes une protéine à activité aconitase très conservée. Ces protéines montrent une forte homologie avec les IRE-BP et devraient avoir un rôle dans des mécanismes de régulation traductionnelle.

**EXEMPLE 7 : ISOLEMENT, SEQUENCAGE ET DETERMINATION DU
DEPART DE TRANSCRIPTION DU GENE ACONITASE
D'ARABIDOPSIS THALIANA :**

Jusqu'à présent, aucun gène codant pour l'aconitase mitochondriale ou pour l'aconitase cytosolique n'a été étudié dans le domaine végétal ou dans le domaine animal.

Deux banques génomiques ont dû être criblées pour pouvoir isoler le gène complet de l'aconitase étudiée.

Dans un premier temps une banque LAMBDA DASH d'Arabidopsis thaliana (écotype C24) (ADN génomique digéré partiellement par l'enzyme de restriction Hind III, puis cloné dans le vecteur DASH au site Hind III) est criblée avec l'ADNc n°1 de 3,2 Kb isolé d'Arabidopsis thaliana.

Parmi les 2.10^5 phages étalés, 6 clones positifs ont été isolés. Ces clones apparaissent identiques après analyse par restriction.

Les analyses par hybridation sur différentes digestions (simples ou doubles) par des enzymes de restriction a permis de montrer que ce clone recouvrait l'extrémité 5' du gène aconitase et était incomplet en 3';

Pour obtenir la partie manquante du gène en 3', les inventeurs ont dû cribler une autre banque génomique.

La banque génomique EMBL 3 d'Arabidopsis thaliana (écotype Columbia) construite après digestion partielle de l'ADN par l'enzyme de restriction Mbo I puis clonage de ces fragments au site BamHI du vecteur EMBL 3. Après étalement de 2.10^5 phages puis hybridation avec le fragment de 0,6 Kb, un clone répondant positivement a été isolé. Un fragment de restriction Hind III de 3,5 Kb hybridant faiblement

avec la sonde de 1,6 Kb et fortement avec la sonde de 0,6 Kb a été cloné dans le vecteur pBS II SK+ puis séquencé jusqu'à ce que l'on retrouve l'extrémité 3' de l'ADNc.

Les hybridations ont permis de compléter la carte de restriction du gène codant pour une aconitase et de vérifier la jonction Hind III entre les deux clones physiques (Figure 14).

Dans le même temps, un Southern a été réalisé pour tenter de voir s'il existe un ou plusieurs gènes codant pour l'aconitase, chez Arabidopsis thaliana, et pour vérifier que la carte de restriction du gène établie à partir des clones phagiques suit bien la carte génomique.

L'ADN génomique de deux écotypes d'Arabidopsis thaliana (Columbia et C24) a été digéré par plusieurs enzymes de restriction. Les fragments obtenus ont été séparés sur un gel d'agarose 0,8 % puis transférés sur une membrane de nylon Hybond N+. Cette membrane a été hybridée avec les fragments EcoRI du clone ADNc n°1 (0,8 ; 1,6 ; 0,6 Kb). La cartographie du locus de ce gène par Southern montre qu'elle suit exactement la carte de restriction obtenue avec les clones phagiques.

Lors des cribles des banques génomiques, il a été utilisé deux écotypes d'Arabidopsis thaliana (Columbia et C24). Sur le Southern, il peut être constaté qu'il n'existe pas de polymorphisme entre C24 et Columbia, au locus étudié, pour les couples enzyme-sonde testés.

L'analyse de ces résultats semble donc indiquer qu'il n'existerait qu'un gène chez Arabidopsis thaliana codant pour l'ADNc isolé.

Les fragments d'ADN génomique suivants ont été clonés dans le vecteur pBS II SK+ avant séquençage :

53

- BamHI - BamHI : environ 7,7 Kb ;
- EcoRI - BamHI : environ 0,6 Kb ;
- EcoRI - EcoRI : environ 2,8 Kb (dont un site EcoRI provenant du site de clonage du vecteur) ;
- BamHI - BamHI : environ 2,0 Kb (dont un site BamHI provenant du site de clonage du vecteur) ;
- BamHI - BamHI : environ 1,3 Kb ;
- Hind III - Hind III : environ 5,5 Kb.

La totalité du fragment Hind III - Hind III de 5,5 Kb a été séquencée. Chaque fragment cloné a été séquencé aux deux extrémités avec les amorces universelles T3 et M13. La stratégie de progression par oligonucléotides a été adoptée pour compléter la séquence.

Il est à noter qu'il existe un site Hind III distant de 0,1 Kb du site Hind III de clonage. La présence de ce site a pu être montrée grâce aux hybridations avec la sonde de 1,6 Kb (la double digestion BamHI-Hind III donne deux fragments de 1,8 Kb et 1,3 Kb alors que la simple digestion BamHI donne un fragment de 2 Kb). La présence de ce site a été confirmée lors du séquençage des extrémités des fragments EcoRI-EcoRI de 2,8 Kb et BamHI-BamHI de 2,0 Kb.

De plus, le fragment de restriction Hind III de 3,5 kb de Lambda 3' EMBL a été cloné dans le vecteur pBS II SK+ puis séquencé jusqu'à ce que l'extrémité 34 de l'ADNc soit retrouvée.

Le gène de l'aconitase d'Arabidopsis thaliana a été séquencé sur une longueur de 6 762 pb (Figure 1).

La comparaison de la séquence du gène de l'aconitase avec l'ADNc correspondant montre que le gène est constitué de 20 exons et de 19 introns.

D'autres gènes montrent également un grand nombre d'exons : le gène de la transferrine humaine est constitué de 17 exons, celui du récepteur de la transferrine de 19 exons.

Il est intéressant de noter la présence du premier intron juste après le codon d'initiation méthionine putatif (position +91). Ceci renforce l'idée que cette enzyme doit être très conservée : l'épissage devant être parfait pour obtenir une protéine fonctionnelle.

L'analyse de la partie 3' non codante montre trois sites de polyadénylation possibles suivant le consensus AATAAA. Ces trois sites sont situés en position 5068, 5199, 5285. Il a été constaté lors de la caractérisation des ADNc (Exemple 5) que ces sites doivent être fonctionnels.

L'extrémité 5' de l'ADNc codant pour l'aconitase a été retrouvée sur la séquence génomique. Cependant, le site d'initiation de la transcription n'était pas connu (la région amont du promoteur est riche en AT, il existe donc plusieurs TATA box putatives). Les résultats de Northern indiquent que le clone ADNc était pratiquement complet (un signal d'hybridation était visualisé à environ 3,2 Kb, identique à la taille de l'ADNc). La présence d'un intron dans la région 5' du gène ne peut cependant pas être exclue.

Aussi, le site d'initiation de transcription a-t-il été cartographié par la technique RACE avec le kit amplifinder RACE CLONTECH (Figure 9). Pour cela, des ARN polyadénylés ont été isolés à partir d'ARN totaux extraits de feuilles. L'intégrité des ARN messagers est contrôlée en Northern en utilisant comme

55

sonde les fragments d'ADN obtenus par digestion EcoRI du clone ADNc n°1.

L'oligonucléotide n° 19 :

GCTGGAAGTCAAGCTCCATGTTTGCCTGCA

positionné en +524 sur la séquence de l'ADNc a servi à synthétiser le premier brin d'ADNc à partir des ARN polyadénylés purifiés.

A l'extrémité 3' du brin d'ADNc nouvellement synthétisé, un oligonucléotide (anchor) du kit 3'NH3 GGAGACTTCCAAGGTCTTAGCTATCACTTAAGCAC 5' est accroché avec la T4 RNA ligase. L'amplification de l'extrémité 5' de l'ADNc se fait par PCR en utilisant un oligonucléotide complémentaire de l'oligonucléotide "anchor" du kit et l'oligonucléotide n° 16 :

CAAGCAAGATCAACAACAGCAGGAACACCA

déterminé à partir de la séquence de l'ADNc positionné en +375. La réaction PCR est effectuée dans les conditions suivantes :

- Dénaturation de l'ADN 5 mn à 94°C ;
- 35 cycles d'amplification :
 - dénaturation 1 mn à 94°C,
 - hybridation 1 mn à 50°C,
 - élongation 2 mn à 72°C ;
- Réaction finale d'élongation 7 mn à 72°C.

Après amplification, aucune bande n'est visualisée sur gel d'agarose coloré au bromure d'éthidium.

Il a alors été réalisé une seconde amplification par PCR à partir de la réaction précédente en prenant

plusieurs couples d'amorces (l'oligonucléotide complémentaire de l'oligonucléotide "anchor" accroché au brin d'ADNc associé avec les oligonucléotides n°16, 9 et 14 respectivement en position +375, +52, +230 désignés à partir de la séquence ADNc). Les séquences des amorces 9 et 14 sont les suivantes :

- n° 14 : CACAGTTACGTATGGCCGA
- n° 9 : GTCGAGTGAAGAAGCAA

Les conditions d'amplification sont celles décrites précédemment. La solution de réaction de la première amplification est précipitée puis reprise dans 10 μ l d'eau stérile. 2 μ l de cette solution serviront de cible pour la deuxième réaction d'amplification par PCR. Le témoin positif du kit attendu à environ 490 pb est visible piste 4. Une réamplification avec l'oligonucléotide n°16 donne de l'aspécificité (piste 2). L'oligonucléotide n°9 semble permettre l'amplification d'une bande d'ADN à environ 200 pb (piste 3). Avec l'oligonucléotide n°14, deux bandes ont été amplifiées (piste 5) d'une taille d'environ 300 et 350 pb. Ces deux bandes ont été clonées dans un vecteur pGEM T sans avoir été préalablement séparées. 35 clones portant une insertion ont été séquencés. La sélection des clones a été réalisée par digestion avec l'enzyme de restriction Bgl II (site de coupure en position +48). Le +1 a été positionné sur la séquence au site comptant le plus grand nombre de clones "cappés" (en l'occurrence 14 sur 35 clones séquencés). Parmi ces 35 clones, 14 s'arrêtent en position +1, 7 en position -2, 4 en position -4, 5 en position -9 et 1 en position +9. Tous ces clones, lors de la séquence montrent un G supplémentaire en 5' qui n'est pas

présent sur la séquence génomique. Il s'agit du capping (coiffe) du messenger (indiquant ainsi avec certitude le site d'initiation de la transcription).

Ainsi, ces résultats suggèrent que le site d'initiation de transcription de ce gène est en fait constitué de plusieurs sites d'initiation (5 ont été déterminés) s'étendant sur une zone de 19 pb (-9 à +9).

Il a été également observé un clone s'arrêtant en position +14, 3 en position +20 et 1 en position +29. Ces clones doivent provenir de l'arrêt prématuré de la reverse transcriptase (aucun G supplémentaire en 5' de ces clones).

Un autre clone a été séquencé et placé en position -49 sur la séquence génomique. Ce clone ne porte pas de G supplémentaire ce qui indique qu'il n'est pas complet. La distance avec le premier site d'initiation est de 50 pb. Ce clone doit appartenir à une deuxième classe de clone amplifiée lors de la PCR et formant la bande d'ADN sur gel d'agarose de plus grande taille. Ce type de résultat pourrait amener à penser qu'il existe un deuxième promoteur plus en amont.

Il est à remarquer que les clones correspondant à la bande haute de l'amplification ne sont pas clonés avec la même efficacité que ceux correspondant à la bande basse.

Bien qu'il ne soit pas possible de conclure quant à la position d'une deuxième site d'initiation de la transcription putatif, il est surprenant de trouver une structure secondaire ressemblant aux IRE (Iron Responsive Element) aux positions -85 et -52 (figure 15).

Dans la partie 5' du gène, un site d'initiation de la transcription a été cartographié avec précision.

En réalité, 5 sites permettent le départ de la transcription dans la région -9 à +9, le site le plus usuel se situerait en +1. Une TATA box putative est située en position -32 (TATAA). Cette position est en accord avec les TATA box décrites qui se situent à 32 ± 7 pb en amont du site d'initiation de la transcription. De multiples sites d'initiation de la transcription seraient présents plus particulièrement pour les gènes de ménage.

De ce site d'initiation de la transcription jusqu'au premier ATG (sans doute le codon d'initiation) en position +91, la séquence est codante, ce qui pourrait faire penser à la présence d'un peptide signal mitochondrial (le cadre de lecture reste ouvert jusqu'au codon TAA en position -30). Mais, dans cette séquence se trouvent un résidu aspartique (en position 25) et un résidu glutamique (en position 14), les résidus acides étant généralement absents dans les peptides signaux mitochondriaux. Il n'existe cependant pas de séquence consensus stricte pour les peptides signaux de mitochondries. Aussi, il est possible que cette séquence constitue bien un signal d'importation dans les mitochondries.

Une telle situation pourrait impliquer la présence d'un autre site d'initiation de la transcription plus en amont et donc également la présence d'un deuxième promoteur.

Une autre situation permettant à un gène de coder pour deux protéines à localisations différentes est possible lorsque l'ARNm porte deux sites d'initiation de la traduction. De tels exemples ont été décrits pour les gènes de levure codant pour les enzymes de synthèse des ARN de transfert de l'histidine et de la valine (NATSOULIS et al, (1986, The HTS1 Gene encodes

both the Cytoplasmic and Mitochondrial Histidine tRNA synthetases of *S. cerevisiae*, Cell 46 : 235-243) ; CHATTON et al, (1988, The yeast VAS1 Gene encodes both Mitochondrial and Cytoplasmic Valyl-tRNA Synthetase, The Journal of Biological Chemistry 263 : 52-57). Ces enzymes existent à la fois dans le cytoplasme et dans l'espace matriciel mitochondrial. Le premier site d'initiation de la traduction permet la traduction d'une protéine avec le peptide signal d'importation dans les mitochondries. Le deuxième site d'initiation de la traduction situé plus en aval donnera une protéine sans peptide d'importation mitochondriale et restera donc à localisation cytoplasmique. Pour le gène de l'aconitase d'*Arabidopsis thaliana*, il n'existe apparemment pas un deuxième site d'initiation de la traduction entre la méthionine en position +91 et le site d'initiation de la transcription (phase de lecture ouverte). Cependant, la présence d'un codon d'initiation non conventionnel ne peut être exclue. En effet, il semble que les codons ACG et AUA puissent jouer le rôle de codon d'initiation de la traduction.

EXEMPLE 8 : EXPRESSION SPATIO-TEMPORELLE DE L'ARN MESSAGER DE L'ACONITASE CHEZ ARABIDOPSIS THALIANA ET BRASSICA NAPUS :

Les ARN totaux ont été extraits de différentes parties des plantes étudiées. Après séparation sur un gel d'agarose dénaturant 1,2 %, les ARN sont transférés sur une membrane de nylon puis hybridés avec les fragments EcoRI du clone ADNC n°1.

Un signal plus ou moins intense est observé dans toutes les parties de la plante, indiquant une expression constitutive du gène de l'aconitase.

Un signal est également obtenu pour les ARN totaux du colza (Brassica napus) indiquant une forte conservation de l'aconitase entre ces deux espèces végétales de la famille des crucifères. Une forte expression de ce gène est détectée sur des jeunes plantules en germination. Or, il apparaît que le cycle du glyoxylate doit être très élevé à ce moment du développement permettant la transformation des lipides en sucres indispensables à la croissance de la plante. Ainsi, l'aconitase étant une enzyme du cycle du glyoxylate, une forte expression de celle-ci à ce stade du développement est en accord avec les besoins métaboliques de la plante. Il est également possible que cette forte expression permette la production de squelettes carbonés nécessaires à plusieurs voies de biosynthèse notamment lors de la synthèse des acides aminés. BROUQUISSE et al, 1987 (supra), avaient mis en évidence une forte activité aconitase cytosolique pouvant jouer cette fonction en association avec une activité isocitrate déshydrogénase NADP⁺ cytosolique, la formation d'équivalents réducteurs NADPH pouvant également être utilisée dans des réactions de biosynthèse. Une forte expression du gène de l'aconitase est également observée dans les fleurs ouvertes alors que les bourgeons en cours de développement ne présentent qu'une faible expression. Pour préciser quel organe floral exprime fortement le gène de l'aconitase, les inventeurs ont travaillé sur le colza (Arabidopsis thaliana étant une plante de petite taille, le prélèvement des différents organes floraux est relativement délicat). Il a alors été observé une forte expression du gène de l'aconitase dans les anthères comparativement à une faible expression dans les pistils. Cette forte activité dans

les anthères pourrait refléter un métabolisme très actif pour la maturation des grains de pollen.

De même, une expression relativement importante est observée dans les siliques immatures. Cette expression pourrait permettre un stockage de l'ARN messager ou de la protéine à activité aconitase dans les graines qui seront utilisés lors de la germination au moment de la reprise des métabolismes. Malgré le bruit de fond important (difficulté d'extraction des ARN de graines du fait d'une présence importante de lipides) un signal est également visible dans les graines sèches. Enfin, une expression relativement faible dans les feuilles et dans les tiges est montrée.

L'ensemble de ces résultats semble indiquer une expression constitutive du gène codant pour une protéine à activité aconitase. Une forte expression lors de la germination peut faire penser à l'activité aconitase cytosolique intervenant dans le cycle du glyoxylate. Il apparaît toutefois que, lors de la mise en place de l'appareil photosynthétique, le cycle du glyoxylate est stoppé, les ARNm ne devant plus être détectés dans les différents tissus de la plante.

Cependant, l'expression observée du gène de l'aconitase dans les feuilles et les tiges pourrait être due à l'activité aconitase cytosolique permettant de produire des squelettes carbonés pour différentes voies de biosynthèse.

Toutefois, si les aconitases cytosoliques et mitochondriales sont codées par le même gène, il pourrait également s'agir du messager de la forme mitochondriale qui est exprimé.

Il est intéressant de noter que toute l'activité de la citrate synthase des graines de ricin est due à l'isoenzyme mitochondriale et que seul l'ARN messager

codant pour la forme mitochondriale serait présente dans les graines (ZEHLER et SCHNARRENBARGER, (1984, Citrate Synthases from Germinating Castor Bean Seeds, I - Purification and properties, Physiologia Plantarum 60 : 1-8)). Ces travaux ont été possibles car les isoenzymes de la citrate synthase mitochondriale et glyoxysomale peuvent être distinguées l'une de l'autre (la citrate synthase glyoxysomale est inhibée par le 5,5' dithiobis (2 nitrobenzoic acide) : DTNB). Les deux formes isoenzymatiques montrent également des points isoélectriques différents respectivement de 5,9 et 9,1 pour les formes mitochondriales et cytosoliques.

Dans le cas de l'aconitase, la situation est plus complexe car non seulement les deux formes isoenzymatiques (cytosolique et mitochondriale) n'ont pas pu être distinguées l'une de l'autre, mais de plus, la forme cytosolique intervient dans le cycle du glyoxylate. Celle-ci permet également de fournir des intermédiaires du métabolisme acide, ces besoins pouvant être permanents et non pas ponctuels comme le cycle du glyoxylate.

Dans cette situation, il est difficile d'étudier l'expression de l'une ou l'autre des formes de l'aconitase car il est également possible, ainsi que les inventeurs en ont émis l'hypothèse, qu'un seul gène code pour ces deux formes.

Pour confirmer les données moléculaires obtenues sur l'organisation de l'extrémité 5' du gène et sur l'expression spatio-temporelle de son ARNm, les inventeurs ont fait appel aux expériences de transgénèse à l'aide du gène rapporteur codant pour la β -glucuronidase (GUS). Seule la partie 5' flanquante du gène a été étudiée.

EXEMPLE 9 : CARACTERISATION D'UN FRAGMENT PROMOTEUR DU GENE ACONITASE

Au vu de la position du site d'initiation de la transcription (déterminée par la technique 5' RACE), le site de restriction Bgl II (position +46), apparaît idéalement placé pour réaliser une fusion transcriptionnelle avec le gène rapporteur GUS. En effet, il est situé en amont du codon de départ potentiel de l'aconitase cytoplasmique. Le fragment Hind III-Bgl II, d'une taille de 1339 pb a été inséré par un clonage orienté dans le vecteur binaire pBi 101.1, sites Hind III - Bam HI (JEFFERSON et al, (1987, GUS Fusions : β -glucuronidase as a sensitive and versatile Gene Marker in Higher Plants, EMBO Journal 6 : 3901-3907)) pour obtenir le vecteur pBios 170.

Ce dérivé particulier du vecteur binaire pBi 101 a été choisi pour obtenir dans le même temps une fusion traductionnelle entre la région amont, potentiellement codante du gène aconitase et la région codante du gène β -glucuronidase (Figure 16).

Ce vecteur binaire a été introduit dans la souche désarmée d'Agrobacterium tumefaciens C58'3 et un clone recombinant a servi à transformer des racines d'Arabidopsis thaliana (VALVEKENS et al, (1988, Agrobacterium Tumefaciens-mediated transformation of Arabidopsis Root Explants using Kanamycin Selection. Proceedings of the National Academy of Sciences of the United States of America 85 : 5536-5540)). Des cals transformés et des plantules régénérées ont été testées pour la présence d'activité β -glucuronidase par le test histochimique. Une coloration bleue a été détectée à la fois sur cals et plantules, indiquant la fonctionnalité du fragment promoteur cloné.

EXEMPLE 10 : CARACTERISATION QUALITATIVE ET QUANTITATIVE DU PROMOTEUR ACONITASE D'ARABIDOPSIS THALIANA :

L'expression dans les feuilles est détectée à la fois dans le mésophylle et dans les tissus conducteurs, l'activité semblant plus forte dans ces derniers.

Dans les racines, l'observation microscopique révèle une activité β -glucuronidase se limitant essentiellement à la zone d'élongation des pointes racinaires.

Lors de l'analyse par hybridation sur des ARN de différents organes et à différents stades, les chercheurs avaient détecté une activité forte dans des fleurs matures d'Arabidopsis. Les résultats sur les organes de reproduction mâle et femelle de colza en développement montraient que cette augmentation intervenait dans la partie mâle. Les observations d'activité β -glucuronidase sur les fleurs transgéniques confirment ces résultats et suggèrent fortement qu'il s'agit d'une augmentation de l'activité transcriptionnelle du promoteur aconitase dans les grains de pollen.

D'autre part, les inventeurs ont pu montrer que le promoteur conduisait à une assez forte activité β -glucuronidase dans l'albumen comparativement aux parties maternelles (téguments, placenta, tissu ovarien).

Cette activité dans l'albumen n'est pas surprenante car il s'agit d'un tissu de réserve fort utile lors de la germination. Bien que les inventeurs soient en présence d'une espèce exalbuminée, ils ont pu visualiser précisément l'expression de ce promoteur aconitase dans ce tissu en cours de

développement. L'embryon immature au stade coeur présente lui aussi une activité β -glucuronidase ; à des stades plus avancés, une expression constitutive est observée dans l'embryon.

Sur une graine mature, cette activité est restreinte à l'embryon et éventuellement à l'albumen résiduel.

Lors de la germination de la graine, l'activité du promoteur aconitase est très forte dans l'embryon et plus particulièrement dans les cotylédons, ces derniers étant les organes de réserve où le cycle de glycosylate est actif.

Des analyses comparatives avec le promoteur CaMV 35S montrent que le promoteur aconitase est environ 20 fois plus faible dans les feuilles âgées (voir tableau II, ci-dessous) :

TABLEAU II:

CONSTRUITS (nom du vecteur)	PLANTES (transformant primaire)	ACTIVITE β -GLUCURONIDASE (nmol 4-MU/mg protéine/minute)
35S - GUS (pBi 121)	40 - 2	10.99
	40 - 4	14.69
Aco-GUS (pBiOS 170)	40 - 6	0.27
	40 - 9	0.69

(dosages effectués selon le protocole de Jefferson et al, 1987, supra)

EXEMPLES DECRIVANT L'ISOLEMENT ET LA CARACTERISATION
DU GENE ACONITASE MAIS

**EXEMPLE 11 : OBTENTION D'UN FRAGMENT D'ADN
COMPLEMENTAIRE D'UN GENE ACONITASE DU MAIS :**

Compte tenu des alignements de séquences protéiques des différentes Aconitases et IRE-BP (cf Figure n° 2), les inventeurs se sont attachés à définir des zones conservées entre toutes ces protéines et si possible identiques entre les 2 aconitases végétales qui avaient préalablement été caractérisées. Ces zones protéiques d'une longueur de 6 à 7 acides aminés devaient à la fois être distantes d'une taille compatible avec la technologie PCR, être dispersées sur toute la longueur de la protéine et surtout être composées d'acides aminés pour lesquels très peu de codons existent, ce dernier point étant impératif afin de concevoir des oligonucléotides dégénérés engendrant un nombre de combinaisons limitées. Aussi, des zones contenant au moins une méthionine ou un tryptophane ont été préférentiellement retenues. Ainsi 6 zones ont été choisies et 6 oligonucléodides dégénérés, 3 sens ou brins codants et 3 antisens ou brins complémentaires ont été synthétisés. Afin de faciliter les étapes de clonage, des bases supplémentaires définissant des sites de restriction ont été rajoutées à l'extrémité 5' de chaque oligonucléotide : site EcoRI pour les oligonucléotides sens (DI, DIV, DV) ; site HindIII pour les oligonucléotides complémentaires (DII, DIII, DVI). Les différents oligonucléotides sont présentés ci-dessous avec la position des acides aminés de la séquence protéique d'Arabidopsis et avec le nombre de

combinaisons correspondantes (chiffres entre parathèses). Voir figure 21.

Conformément aux données obtenues dans l'exemple 2 et précisées dans la figure 11(b), l'oligonucléotide dégénéré DI a été conçu en préférant, pour l'acide aminé en position 32, une Alanine à la Sérine trouvée chez Arabidopsis. En effet, la séquence de la pomme de terre et du melon présentaient cet amino-acid.

Les oligonucléotides DII, DIII et DVI ont séparément servi à amorcer des réactions de transcription réverse sur des ARN poly A⁺ isolés et purifiés à partir de plantules de maïs, étiolées et âgées de 6 jours. Ces réactions ont été réalisées avec le kit Reverse Transcription System de la société Promega et ont conduit à l'obtention de trois ADN complémentaires. Sur ces derniers, plusieurs réactions PCR ont été réalisées avec différentes combinaisons d'oligonucléotides DI/DII ; DI/DIII ; DI/DIV ; DV/DVI et cela dans les conditions suivantes : 1 microlitre d'ADN complémentaire est ajouté à 50 microlitres de mélange PCR composé de dNTP 0,25 mM, 1,5 mM de MgCl₂, 3 unités de Taq Polymérase (Promega), 50 mM KCl, 10 mM Tris-HCl, 0,2 microm de chaque oligonucléotide. La réaction est réalisée dans un DNA Thermal Cycler (Perkin Elmer Cetus Version 2.2) selon les étapes suivantes : 3 min de dénaturation à 95°C suivies de 40 cycles composés de 30 secondes de dénaturation à 95°C, 45 secondes d'hybridation à 55°C et d'une minute et 30 secondes d'extension à 72°C.

Seulement une réaction d'amplification (amorces DIII et DIV) a donné un produit qui, de plus, présentait, après électrophorèse en gel d'agarose, une taille compatible avec le fragment théorique attendu (100 pb environ). Ce produit a été introduit par ligation dans le vecteur pGEM-T de Promega selon les

recommandations du fournisseur (pGEM-T Vector System I) et transformé dans la souche d'*Escherichia coli* XLI Blue obtenue auprès de la société Stratagène.

Divers clones ont été obtenus et le fragment inséré de l'un d'entre eux a été séquencé selon le protocole du Kit de Promega, fmol DNA Sequencing System. La séquence qui a été déterminée démontrait que ce fragment était hautement homologue avec la séquence d'*Arabidopsis thaliana*. Une homologie nucléique de 77,5 % et protéique de 87,5 % était détectée (figure n° 17).

EXEMPLE 12 : ISOLEMENT ET DETERMINATION DE LA SEQUENCE D'UN GENE ACONITASE DU MAIS

Une banque génomique EMBL 3 de maïs construite après digestion partielle par l'enzyme MboI de l'ADN isolé de plantules (stade 2 feuilles) de la lignée B73 a été obtenue (Clontech référencée FL 1032 D). Après étalement de $2 \cdot 10^6$ phages puis hybridation avec le fragment d'ADN complémentaire de maïs de 100 bp identifié dans l'exemple 11 (figure 17), 9 clones génomiques ont été isolés. Quatre d'entre eux se sont révélés identiques après analyse par restriction.

Sur un des clones, le clone T1, un fragment SalI de 6,2 Kb hybridant avec le fragment d'ADNc a été sous cloné dans le vecteur pBS II SK+. Ce fragment a été séquencé initialement aux extrémités ainsi qu'à partir de 2 oligonucléotides spécifiques du fragment d'ADN complémentaire puis, ensuite, par la stratégie de progression par oligonucléotides. Les premières comparaisons avec l'ADN complémentaire d'*Arabidopsis* ont montré que ce clone ne contenait pas l'extrémité 5' du gène aconitase de maïs.

Afin d'essayer d'obtenir la séquence de l'extrémité 5', les inventeurs ont entrepris de séquencer par la technique de progression par oligonucléotides la zone amont sur les autres clones génomiques. Seulement 2 d'entre eux, R1 et J1 présentaient une zone 5' du gène plus étendue que le clone T1 ; avec toutefois J1 qui ne contenait que 1060 pb de plus que le clone T1.

R1 contient une insertion d'environ 15 kb d'ADN génomique de maïs, comprenant notamment un fragment SalI-HindIII d'environ 6800 pb recouvrant l'extrémité 5' du gène aconitase, et s'étendant de la position -4000 à la position 2841 (figure 18). Ce fragment est purifiable par double digestion HindIII, SalI ; SalI étant un site de restriction unique présent sur le vecteur et contigu au site d'insertion du fragment génomique, à savoir BamHI, compatible avec l'enzyme MboI.

C'est au total 6036 paires de bases de la séquence d'un gène aconitase de maïs qui ont pu être déterminées du clone R1 (Figure 18) avec une seule incertitude de quelques bases à la position 4635 ; cette région correspond à un intron. Etant donné que les inventeurs ne disposaient pas de clone d'ADN complémentaire synthétisé à partir de l'extrémité polyA+ de ce gène aconitase de maïs, les régions codantes ont été définies par recherche d'homologies avec les séquences codantes d'Arabidopsis et celle de l'ADN complémentaire incomplet du melon. Ceci a permis de démontrer qu'à partir des positions 530-540, la zone codante est très homologue avec l'exon 2 d'Arabidopsis et que, par la suite, absolument tous les exons d'Arabidopsis ont été retrouvés, attestant ainsi que l'organisation génétique du gène aconitase est très conservée entre ces 2 espèces. Comme les

inventeurs n'ont pas pu déterminer l'extrémité 5' de l'exon 2 et qu'aucune homologie nette avec l'exon 1 n'est définissable dans les 500 premières bases de cette séquence, les exons ont été numérotés en fonction des homologues Arabidopsis ; ils s'étendent donc des positions :

- position exacte à définir à 623 pour l'exon 2 ;
- 1122 à 1282 pour l'exon 3 ;
- 1360 à 1452 pour l'exon 4 ;
- 1664 à 1839 pour l'exon 5 ;
- 1932 à 2093 pour l'exon 6 ;
- 2335 à 2479 pour l'exon 7 ;
- 2572 à 2714 pour l'exon 8 ;
- 3074 à 3130 pour l'exon 9 ;
- 3229 à 3308 pour l'exon 10 ;
- 3407 à 3482 pour l'exon 11 ;
- 3569 à 3754 pour l'exon 12 ;
- 3842 à 3909 pour l'exon 13 ;
- 4030 à 4148 pour l'exon 14 ;
- 4241 à 4362 pour l'exon 15 ;
- 4474 à 4572 pour l'exon 16 ;
- 4686 à 5165 pour l'exon 17 ;
- 5239 à 5337 pour l'exon 18 ;
- 5423 à 5665 pour l'exon 19 ;
- 5764 à 5835 pour l'exon 20 (Partie codante uniquement).

Des signaux potentiels de polyadénylation sont présents aux positions 5069 et 5081. Seul l'exon n° 17 a une taille différente de celui d'Arabidopsis (+ 3 paires de bases) et c'est aussi celui qui code pour la zone protéique présentant le plus de divergence malgré une séquence nucléique assez conservée. Ceci

s'explique par l'addition d'un T à la position 5021 séquence du maïs (ou la délétion d'une base après la position 3047 sur la séquence d'Arabidopsis).

Il est assez remarquable de retrouver tous les introns situés à des positions identiques sur ces 2 gènes. Le tableau III présente la taille des exons et introns des deux gènes Aconitase étudiés. La taille des introns qui est en moyenne supérieure chez le gène du maïs diffère grandement d'un intron à l'autre entre ces 2 gènes : un intron relativement long chez Arabidopsis (intron 17 = 340 pb) peut être court chez le maïs (= 73) et réciproquement (intron 8 d'Arabidopsis = 91 et = 359 chez le maïs).

La figure n° 19 montre l'alignement des aconitases Arabidopsis et maïs attestant de la très grande conservation de ces enzymes. Sur la figure n° 20, c'est un alignement entre les 3 aconitases végétales décrites dans cette invention (ZMACO = Aconitase maïs ; ATACO = Aconitase Arabidopsis ; CMACO = Aconitase melon) présentant les homologies sur la partie commune aux trois séquences.

N°	EXON		INTRON	
	ARABIDOPSIS	MAIS	ARABIDOPSIS	MAIS
1	92	?	152	?
2	105	?	107	498
3	161	161	104	77
4	93	93	91	201
5	186	186	84	92
6	162	162	122	241
7	145	145	114	92
8	143	143	91	359
9	57	57	100	96
10	80	80	79	98
11	76	76	241	87
12	186	186	82	87
13	68	68	85	120
14	119	119	86	92
15	122	122	89	109
16	99	99	89	?
17	477	480	340	73
18	99	99	68	84
19	243	243	107	98
20	67	72		

Tableau III: Tailles des exons et introns des gènes Aconitase d'*Arabidopsis thaliana* et de Maïs

REVENDEICATIONS

1. Protéine ayant une activité enzymatique d'aconitase, caractérisée en ce qu'elle comporte, soit

i) la séquence en acides aminés illustrée dans la figure 1 ou une séquence présentant au moins 75 % d'homologie avec celle-ci, soit

ii) une variante de la séquence i) comportant une délétion de jusqu'à 40 acides aminés environ au niveau de l'extrémité NH₂.

2. Enchaînement d'acides aminés comprenant au moins un fragment de la protéine selon la revendication 1, ledit fragment consistant en :

a) une séquence ayant une longueur comprise entre 6 et 40 acides aminés, ladite séquence présentant, sur toute sa longueur, au moins 90% d'homologie avec la partie correspondante de la séquence en acides aminés de la figure 1, ou présentant au moins 60% d'homologie avec l'une des séquences suivantes :

I : DTVSMIEAYLRANKMFVDYSEPEskTVYSSCLELNLEDVE

II : PLKEMKADWHSCLDNRV

III : AVPKEAQSKAVEFNFNNGTTAQLR

IV : KGMTMSPPG

V : AVVMMRLWREHFANIRIVNKHLKGEVG

VI : NVSEIKPGQDVTVVTTNN

VII : les acides aminés 1 à 40 de l'extrémité NH₂ de la protéine illustrée dans la figure 1.

ou

b) une séquence ayant une longueur supérieure à 40 acides aminés, ladite séquence présentant, sur

toute sa longueur, au moins 75% d'homologie avec la partie correspondante de la séquence en acides aminés de la figure 1.

3. Enchaînement d'acides aminés selon la revendication 2 caractérisé en ce qu'il comprend une séquence ayant une longueur comprise entre 40 et 100 acides aminés et présente, sur toute sa longueur, au moins 80% d'homologie avec la partie correspondante de la séquence en acides aminés de la figure 1.

4. Séquence d'acide nucléique caractérisée en ce qu'elle comporte soit :

i) une séquence comprenant à la fois une partie 5' non-transcrite, une partie transcrite, éventuellement avec des introns, et une partie 3' non transcrite, ladite séquence donnant lieu, après transcription, épissage et traduction à la protéine selon la revendication 1, soit

ii) le transcrit de la séquence i) ou son équivalent ADNc, soit

iii) une séquence complémentaire de la séquence i) ou ii), soit

iv) une séquence ayant une longueur d'au moins 25 bases et étant capable de s'hybrider, sur toute sa longueur, avec les séquences i), ii) ou iii) dans des conditions stringentes ou moyennement stringentes, soit

v) un fragment consistant en au moins 25 nucléotides consécutives de l'une quelconque des séquences i), ii), ou iii), ou en au moins 10 nucléotides consécutives lorsqu'il s'agit d'un fragment d'un intron de la séquence i).

5. Séquence d'acide nucléique selon la revendication 4 caractérisée en ce qu'elle comporte la séquence génomique illustrée dans la figure 1 ou celle

illustrée dans la figure 18, ou un fragment de l'une de ces séquences.

6. Séquence d'acide nucléique selon la revendication 5 caractérisée en ce qu'elle comporte au moins une partie d'une région non-codante ou d'un intron de la séquence génomique.

7. Séquence d'acide nucléique caractérisée en ce qu'elle comporte au moins un intron du gène d'aconitase végétale, ou une partie d'un tel intron, ladite partie ayant au moins 10 nucléotides et de préférence au moins 30 nucléotides.

8. Séquence d'acide nucléique selon la revendication 6 caractérisée en ce qu'elle présente une activité promotrice et se trouve au sein du fragment HindIII(-1293)-BglII(+45) de la séquence génomique de la figure 1.

9. Séquence d'acide nucléique selon la revendication 4, capable d'inhiber le gène de l'aconitase caractérisée en ce qu'elle comporte un ou plusieurs fragments ayant une longueur d'au moins 6 nucléotides, le(s)dit(s) fragment(s) étant complémentaire(s) d'au moins une partie du transcrit du gène de l'aconitase.

10. Couple de séquences d'acide nucléique selon la revendication 4 caractérisée en ce que chaque séquence a une longueur comprise entre 15 et 300 nucléotides et est capable de jouer le rôle d'amorce dans une réaction d'amplification d'acide nucléique.

11. Plasmide caractérisé en ce qu'il contient au moins l'une des séquence d'acide nucléique selon les revendications 5 à 10.

12. Gène chimérique capable d'être exprimé chez les plantes caractérisé en ce qu'il comporte :

i) une région promotrice selon la revendication 8, et

ii) une séquence hétérologue transcrite, codante ou non, placée sous le contrôle dudit promoteur.

13. Gène chimérique capable d'être exprimé chez les plantes caractérisé en ce qu'il comporte :

i) un promoteur fonctionnel chez les plantes autre que celui naturellement associé avec le gène de l'aconitase, et

ii) une séquence codant pour la protéine selon la revendication 1 ou pour un fragment protéique selon la revendication 2.

14. Cellules végétales transformées de manière stable par une séquence d'acide nucléique selon l'une quelconque des revendications 4 à 13.

15. Plantes transgéniques transformées de manière stable par une séquence d'acide nucléique selon l'une quelconque des revendications 4 à 13.

16. Anticorps monoclonaux ou polyclonaux capables de reconnaître, de préférence de manière spécifique, la protéine selon la revendication 1 ou le fragment protéique selon la revendication 2.

17. Procédé pour la modification du métabolisme d'une plante, caractérisé par l'introduction dans la plante d'un insert génétique comportant une séquence d'acide nucléique selon la revendication 4.

18. Procédé selon la revendication 17 caractérisé en ce que la séquence d'acide nucléique comprend un promoteur fonctionnel chez les plantes, et sous le contrôle de ce promoteur, une séquence codant pour la protéine selon la revendication 1, ou pour un fragment protéique selon la revendication 2, l'expression de cette séquence conduisant à une surproduction d'aconitase.

19. Procédé selon la revendication 18 caractérisé en ce que la surproduction d'aconitase conduit à une surproduction d'au moins l'un des acides organiques du

cycle de Krebs ou du cycle glyoxylate, notamment le citrate, l'isocitrate, le succinate, le glyoxylate, le malate, l'oxaloacétate, l' α -cetoglutarate, le fumarate.

20. Procédé selon la revendication 19 caractérisé en ce que le promoteur est soit celui naturellement associé, dans la plante, avec le gène de l'aconitase, soit un promoteur hétérologue au gène de l'aconitase.

21. Procédé selon la revendication 17 caractérisé en ce que l'insert génétique comprend un promoteur fonctionnel, et sous le contrôle de ce promoteur, une séquence inhibitrice selon la revendication 9, l'expression de la séquence conduisant à une inhibition de l'expression du gène de l'aconitase.

22. Procédé selon la revendication 20 caractérisé en ce que l'inhibition de l'expression du gène de l'aconitase conduit à une surproduction de l'Acétyl coenzyme-A et par conséquent, à une réorientation du métabolisme ou du catabolisme polysaccharidique, lipidique et azoté.

23. Procédé pour obtenir l'expression constitutive, dans une plante, d'une séquence codante, l'expression étant particulièrement forte lors de la germination des graines et lors de la maturation de la graine et du pollen, caractérisé par l'introduction dans le génome de la plante, d'un gène chimérique selon la revendication 12.

24. Gène chimérique capable d'être exprimé chez les plantes caractérisé en ce qu'il comporte :

- i) un promoteur fonctionnel chez les plantes,
- ii) une séquence transcrite, codante ou non, placée sous le contrôle dudit promoteur,
- iii) un ou plusieurs introns du gène de l'aconitase végétale, selon la revendication 6 ou 7.

FIGURE 1(i)

HindIII
aagctttcggtccatattaaaataaacattttttttttttatataacttctttctttattgatattag-1225
 cttaaaagatcggtataaaacaaacttatatacgatcaaacaaactaaaaagagaggttctttatcaaaaaaa-1153
 acaaaactaaaaagagacgataaatgaaagagataacaatgatatgtaattgtccttacataaaaaaaaata-1081
 ttgagccacgccaaatctaaaacgatttcgttaaacatatatagttctcgatatttatcaacaaagacgatg-1009
 taattcgtaatctatatacgatcatcttcttaatttttttaaaaaatataaaaagggtgcttcagattcaaggg -937
 actgaaactagtggaggtttcttaaaagcagatatatacactgaccacgttcagggttaatagttagtagattt -865
 ttagggtggatagtaatatattgatggacaatatattaaatacagactataatatttaagagacttgaagtgtt -793
 cctaagagtaaatttaattatacatttttaatacatggcctgaaaaactaaacaaggaggggtagaataattaa -721
 tcatctcagcagctttccatctgacacaccattcagctaaacgacattgtttaatctctaaccaaaactggga -649
 tagaccgaactgggtacaaagatgggttaaaactggctaggcccaaattaatcagaacataaaatttcgaacgg -577
 agcgacccattttccaaaataaaaaaatgtatttttcatatatatatatatttttttttttaaatcggttgcatt -505
 gcaggttggtggcgatatttttttggtttgtcacaggtggtggttaataaaacgacgatcattgaatatattttt -433
 ggtgtcttgatatatgtttgtgtcattgtgtgtatttttttttttggtgtgtgtgtaacgacgatcattcttgta -361
 tatgtttgtcagaaattaagggttttttttaataatcaatcaatgccatgtgaaattgagaaaccttgaagaaaatt -289
 ggattacataaaaactagatactatagagaagaggcacatatcagattggtggattataagtttggtgaaaata -217
 aaatcatgattgaacagatatattaattcacttgatacaaaaaaagcaaagaatggatattcacaaaaaaaag -145
 aaagttggataattaagaaaaatagaatataaaaataaaaaataaatcaaacagccatcagttctgacgaaaa -73
 gagccaccacggatctgtcagattttatttccttcttaca tataa aaaaaatctgaatcggtcaggctcgt -1
Bgl2
 GTA GCC TCA CCA ACT CCG CTT ATA AAC TTT CTC TCT TCT GAA AAC AGA TCT CTC 54
 Val Ala Ser Pro Thr Pro Leu Ile Asn Phe Leu Ser Ser Glu Asn Arg Ser Leu
 TAC TTT GCT TCT TCA CTC GAC TTG TAT CTA TCA TCC ATG G gtatcgattttctctgttt 112
 Tyr Phe Ala Ser Ser Leu Asp Leu Tyr Leu Ser Ser Met
 acgatccaaaatgattcatgtttcttattttcttcttctcgcgatgctgattttccagctatttttcggtttt 184
 tttgaaatctggttctgtatttttgatttgctgtgtggatctctgtttctgatgtgttttcagCT TCC GAG 254
 Ala Ser Glu
EcoRI
 AAT CCT TTC CGA AGC ATA TTG AAG GCG TTA GAG AAG CCT GAT GGT GGT GAA TTC 308
 Asn Pro Phe Arg Ser Ile Leu Lys Ala Leu Glu Lys Pro Asp Gly Gly Glu Phe
 GGT AAC TAC TAC AGC TTA CCT GCT TTG AAC GAT CCC AGG ATC G gtgagtttcctttt 365
 Gly Asn Tyr Tyr Ser Leu Pro Ala Leu Asn Asp Pro Arg Ile
 atctctgatagatattataatattttgttagaagacgttggttatgatctcggttggtgagagttttatgagag 437
 aaatgttggtggtggatgacagAT AAA CTA CCT TAT TCC ATT AGG ATA CTT CTT GAA TCG 496
 Asp Lys Leu Pro Tyr Ser Ile Arg Ile Leu Leu Glu Ser
 GCC ATA CGT AAC TGT GAT GAG TTC CAA GTT AAG AGC AAA GAT GTT GAG AAG ATT 550
 Ala Ile Arg Asn Cys Asp Glu Phe Gln Val Lys Ser Lys Asp Val Glu Lys Ile

2/36

FIGURE 1 (ii)

CTT GAT TGG GAG AAT ACT TCT CCC AAG CAG GTT GAG ATT CCG TTC AAG CCT GCT 604
 Leu Asp Trp Glu Asn Thr Ser Pro Lys Gln Val Glu Ile Pro Phe Lys Pro Ala
 CGG GTT CTT CTT CAG gtagacaatgattggttgctatacacattcttgatgacataatgatgtttc 670
 Arg Val Leu Leu Gln
 tgatttctctgaaaatgctagaatgtgatatcattttttctatttggttttcagGAC TTT ACT GGT GTT 738
 Asp Phe Thr Gly Val
 CCT GCT GTT GTT GAT CTT GCT TGC ATG AGA GAT GCC ATG AAT AAT CTC GGT GGT 792
 Pro Ala Val Val Asp Leu Ala Cys Met Arg Asp Ala Met Asn Asn Leu Gly Gly
 GAT TCT AAT AAA ATT AAT CCG CTG gttagtttctctcgctttttactgagattggtaaatgatt 855
 Asp Ser Asn Lys Ile Asn Pro Leu
 tgtgttcttttggtatgctataaggtttgatacaaggccttgataattgcagGTC CCT GTA GAT CTT 922
 Val Pro Val Asp Leu
 GTC ATT GAC TAC TCC GTT CAG GTG GAT GTG GCG AGA TCA GAG AAC GCA GTG CAG 976
 Val Ile Asp Tyr Ser Val Gln Val Asp Val Ala Arg Ser Glu Asn Ala Val Gln
 GCA AAC ATG GAG CTT GAG TTC CAG CGT AAC AAG GAA AGA TTT GCT TTT CTT AAG 1030
 Ala Asn Met Glu Leu Glu Phe Gln Arg Asn Lys Glu Arg Phe Ala Phe Leu Lys
 BamHI
 TGG GGA TCC AAC GCC TTT CAC AAC ATG CTT GTC GTA CCT CCT GGA TCT GGA ATA 1084
 Trp Gly Ser Asn Ala Phe His Asn Met Leu Val Val Pro Pro Gly Ser Gly Ile
 GTT CAT CAA gtaagtagaggattcagggacaacggtattgcttgaaacttaaaagtatttttatgcta 1152
 Val His Gln
 aattctataattttgcaattgacagGTC AAC CTA GAA TAC CTT GCC AGA GTT GTT TTC AAC 1213
 Val Asn Leu Glu Tyr Leu Ala Arg Val Val Phe Asn
 ACA AAT GGA CTT CTT TAC CCA GAC AGT GTT GTT GGC ACA GAC TCT CAC ACC ACT 1267
 Thr Asn Gly Leu Leu Tyr Pro Asp Ser Val Val Gly Thr Asp Ser His Thr Thr
 ATG ATT GAT GGA CTG GGT GTT GCT GGA TGG GGA GTT GGC GGT ATA GAA GCG GAA 1321
 Met Ile Asp Gly Leu Gly Val Ala Gly Trp Gly Val Gly Gly Ile Glu Ala Glu
 CGA CCG ATG CTT GGT CAG gtaaatcagatgctctttcggtcgagaactcatagctaactgtctaa 1386
 Arg Pro Met Leu Gly Gln
 cctctttctgaaactttggttctaaactacatcctttcaataattctaatagactttccatctcttttatg 1458
 cagCCA ATG AGC ATG GTC CTA CCC GGT GTT GTG GGT TTC AAG CTA ACG GGA AAG 1512
 Pro Met Ser Met Val Leu Pro Gly Val Val Gly Phe Lys Leu Thr Gly Lys
 TTA AGA GAT GGA ATG ACA GCT ACT GAT TTG GTC TTA ACA GTG ACT CAG ATG TTG 1566
 Leu Arg Asp Gly Met Thr Ala Thr Asp Leu Val Leu Thr Val Thr Gln Met Leu
 EcoRI
 AGG AAA CAT GGA GTA GTT GGA AAG TTT GTT GAA TTC CAC G gtacgttgataaaaagatt 1624
 Arg Lys His Gly Val Val Gly Lys Phe Val Glu Phe His
 tgatatcatattttcatatatgtctcttatttacttacatagtagcttaccgagtttaggctagctcacatatg 1696
 catcgactgctttttcatatacagGG GAA GGG ATG AGA GAA TTG TCT TTA GCT GAC CGT 1755
 Gly Glu Gly Met Arg Glu Leu Ser Leu Ala Asp Arg
 GCT ACA ATT GCC AAT ATG TCT CCT GAG TAC GGT GCG ACC ATG GGA TTC TTC CCA 1809
 Ala Thr Ile Ala Asn Met Ser Pro Glu Tyr Gly Ala Thr Met Gly Phe Phe Pro
 GTC GAT CAT GTC ACT TTG CAG TAT CTA AGG TTG ACA GGC AGG AGC GAT GAC ACT 1863
 Val Asp His Val Thr Leu Gln Tyr Leu Arg Leu Thr Gly Arg Ser Asp Asp Thr

FIGURE 1 (iii)

gtaagaaaatgactcacatttcgggaaaatcttcattatcttagtggttaaagactttcataatgggtattga 1935
 tttgtgaattggattgcagGTC TCC ATG ATA GAG GCG TAT TTA CGA GCA AAC AAG ATG 1993
 Val Ser Met Ile Glu Ala Tyr Leu Arg Ala Asn Lys Met
 TTT GTG GAT TAC AGT GAG gtaaacccttgggcatccttaatacttcatttcttatgcctattgtca 2058
 Phe Val Asp Tyr Ser Glu
 tatatcatatttgcataacatatgcttcacttgggaacttgttgcacatctcagCCG GAG AGT AAG ACA 2126
 Pro Glu Ser Lys Thr
 GTT TAT TCC TCA TGT CTG GAA TTG AAT CTC GAG GAT GTG GAA CCT TGT GTT TCT 2180
 Val Tyr Ser Ser Cys Leu Glu Leu Asn Leu Glu Asp Val Glu Pro Cys Val Ser
 GGT CCC AAG AG gtactattagtgaaatgattcgacagaaaattatcttaggatccattctacttacaatg 2248
 Gly Pro Lys BamHI
 tctattctctgtgcacaatacagG CCT CAT GAT CGT GTT CCT TTG AAG GAA ATG AAA GCG 2308
 Arg Pro His Asp Arg Val Pro Leu Lys Glu Met Lys Ala
 GAC TGG CAT TCT TGC TTG GAC AAT AGA GTA GGA TTC AAG gttagtactctcctttctg 2366
 Asp Trp His Ser Cys Leu Asp Asn Arg Val Gly Phe Lys
 ttgagaagatgtgaatgcttgcagtcgccgagagcttaattttagtagaaattatgacacctctagctattggtt 2438
 actgtaattttgaaagaatggtacgtattcgagctctttaaagcgggtccaataaactaacttgtagcttctgc 2510
 aacacttctgaatttataacctacttccattttttcaaaaatagtggttagtgatactacttatccatgatt 2582
 ttccagGGT TTC GCT GTA CCT AAA GAA GCA CAG AGT AAG GCT GTA GAG TTC AAT 2636
 Gly Phe Ala Val Pro Lys Glu Ala Gln Ser Lys Ala Val Glu Phe Asn
 TTT AAC GGG ACC ACA GCA CAG CTT AGA CAT GGA GAT GTT GTT ATA GCA GCA ATC 2690
 Phe Asn Gly Thr Thr Ala Gln Leu Arg His Gly Asp Val Val Ile Ala Ala Ile
 ACC AGT TGC ACA AAT ACT TCA AAC CCT AGT GTA ATG CTT GGC GCT GCC TTA GTT 2744
 Thr Ser Cys Thr Asn Thr Ser Asn Pro Ser Val Met Leu Gly Ala Ala Leu Val
 GCA AAA AAG GCC TGC GAC CTA GGA CTG GAG gtttgactacttgtgtgatagtgatttaata 2805
 Ala Lys Lys Ala Cys Asp Leu Gly Leu Glu
 tctttatgtgcaatggaaataatagaagtgtttgaaatttctcaattatagGTT AAG CCA TGG ATC 2871
 Val Lys Pro Trp Ile
 AAA ACT AGT CTT GCT CCA GGC TCT GGA GTT GTA ACA AAG TAC TTG GCA AAG AG 2924
 Lys Thr Ser Leu Ala Pro Gly Ser Gly Val Val Thr Lys Tyr Leu Ala Lys
PstI
 gtctgctgcagatacacccgttgaaatgttttattttatcttctccttgcattgtgtgtggtcctgagttt 2996
 tcaagtcttgcagT GGC TTG CAG AAG TAC TTG AAT CAG CTC GGC TTC AGT ATC GTT 3052
 Ser Gly Leu Gln Lys Tyr Leu Asn Gln Leu Gly Phe Ser Ile Val
 GGT TAT GGG TGC ACC ACA TGC ATT GGA AAC TCG GGG GAT ATC CAT GAA GCT GTG 3106
 Gly Tyr Gly Cys Thr Thr Cys Ile Gly Asn Ser Gly Asp Ile His Glu Ala Val
 GCT TCA GCA ATA GTT GAT AAT G gtaattccttaagtagttaatgaagtaatgcttcttcacttg 3170
 Ala Ser Ala Ile Val Asp Asn
 aattttcgaagggtcaaaactaaaagatttctcatatatgtacaagAC TTG GTG GCA TCC GCT GTG 3234
 Asp Leu Val Ala Ser Ala Val
 TTG TCT GGG AAC AGA AAT TTT GAG GGA CGT GTT CAC CCG TTA ACA AGA GCT AAC 3288
 Leu Ser Gly Asn Arg Asn Phe Glu Gly Arg Val His Pro Leu Thr Arg Ala Asn

FIGURE 1 (iv)

TAT CTA GCT TCC CCA CCG CTT GTT GTA GCC TAT GCT CTG GCT GGG ACT gtatgtt 3343
 Tyr Leu Ala Ser Pro Pro Leu Val Val Ala Tyr Ala Leu Ala Gly Thr
 aatgcaactctaaagccatttcacaacattcaagcattcttggttatatgaagcacagaatcattcacatgat 3415
 attcttgtagGTT GAC ATT GAT TTT GAG ACA CAG CCC ATT GGA ACT GGG AAA GAT 3470
 Val Asp Ile Asp Phe Glu Thr Gln Pro Ile Gly Thr Gly Lys Asp
 GGA AAA CAG ATA TTT TTC AGG GAC ATT TGG CCC TCT AAC AAA GAA GTT GCT GAG 3524
 Gly Lys Gln Ile Phe Phe Arg Asp Ile Trp Pro Ser Asn Lys Glu Val Ala Glu
 gtaaataatatatgggctacttgggtaatacgtaaacaatgagatattataagtggaggggatgccaaattct 3596
 ctgtttcgattttttcagGTT GTT CAA TCT AGT GTC CTT CCT GAT ATG TTC AAA GCT ACA 3655
 Val Val Gln Ser Ser Val Leu Pro Asp Met Phe Lys Ala Thr
 TAT GAA GCA ATC ACC AAA GGA AAT TCC ATG TGG AAT CAG TTA TCT GTG GCG TCA 3709
 Tyr Glu Ala Ile Thr Lys Gly Asn Ser Met Trp Asn Gln Leu Ser Val Ala Ser
 GGT ACT CTC TAT GAG TGG GAC CCG AAA TCA ACT TAC ATT CAC GAG CCG CCT TAT 3763
 Gly Thr Leu Tyr Glu Trp Asp Pro Lys Ser Thr Tyr Ile His Glu Pro Pro Tyr
 TTC AAG GGC ATG ACC ATG TCT CCA CCC GGT CCA CAT GGT GTG AAA GAC GCA TAC 3817
 Phe Lys Gly Met Thr Met Ser Pro Pro Gly Pro His Gly Val Lys Asp Ala Tyr
 TGT TTA CTC AAT TTT GGA GAC AGT ATT ACC ACT GAT CAC ATC TCA CCA GCT GGT 3871
 Cys Leu Leu Asn Phe Gly Asp Ser Ile Thr Thr Asp His Ile Ser Pro Ala Gly
 AGC ATC CAC AAG GAC AGT CCT GCG GCT AAG TAC TTG ATG GAA CGA GGT GTG GAT 3925
 Ser Ile His Lys Asp Ser Pro Ala Ala Lys Tyr Leu Met Glu Arg Gly Val Asp
 AGA AGA GAC TTC AAC TCA TAC GGA GTC GCC GTG GTA ATG ATG AGA TTA TGG CGA 3979
 Arg Arg Asp Phe Asn Ser Tyr Gly Val Ala Val Val Met Met Arg Leu Trp Arg
 GAG CAC TTT GCA AAT ATC CGT ATT GTC AAC AAA CAC TTG AAA GGA GAA GTT GGT 4033
 Glu His Phe Ala Asn Ile Arg Ile Val Asn Lys His Leu Lys Gly Glu Val Gly
 CCC AAA ACA GTT CAC ATT CCC ACT GGA GAG AAG CTT TCT GTT TTC GAT GCT GCC 4087
 Pro Lys Thr Val His Ile Pro Thr Gly Glu Lys Leu Ser Val Phe Asp Ala Ala
 ATG gtataatctcactttttttttttatccttttgaaaacatcaaattcaagctctttgggtcatgaggttg 4157
 Met
 HindIII
 gggtaggattacgataaagctttgaataagcactacagctctactcttagctataggtctactgtatctagcc 4229
 taggactgaatgtagtagattagagctgttgtaggtgtgtttatagaacagttagtgaaaggtgaggttaata 4301
 ataatgtctcagcagattgtcttgggggtgaaattttgtgtgttgattctctgatttggtcagttggaatt 4373
 aagaatagcagttcatagctttgttttaacagattattctctcctcgtgctgtacagAAA TAT AGG 4440
 Lys Tyr Arg
 AAC GAG GGA CGC GAC ACA ATC ATT TTG GCT GGT GCT GAA TAC GGT AGT GGA AGT 4494
 Asn Glu Gly Arg Asp Thr Ile Ile Leu Ala Gly Ala Glu Tyr Gly Ser Gly Ser
 TCT CGT GAT TGG GCT GCC AAG GGT CCA ATG CTT CTG gtacattttctgtatttaacaaa 4553
 Ser Arg Asp Trp Ala Ala Lys Gly Pro Met Leu Leu
 cacgttcttacactttatttttatgcggtttcttttaaaaaattgtgacaacagGGT GTG AAA GCT GTG 4619
 Gly Val Lys Ala Val
 EcoRI
 ATT TCA AAG AGC TTC GAG CGA ATT CAC CGA AGC AAT TTG GTG GGA ATG GGA ATC 4673
 Ile Ser Lys Ser Phe Glu Arg Ile His Arg Ser Asn Leu Val Gly Met Gly Ile

FIGURE 1(v)

ATA CCT TTG TGC TTC AAG GCG GGA GAA GAT GCT GAG ACC CTT GGC CTA ACG GGT 4727
 Ile Pro Leu Cys Phe Lys Ala Gly Glu Asp Ala Glu Thr Leu Gly Leu Thr Gly
 CAG GAG CTT TAC ACC ATT GAG CTC CCA AAC AAT GTT AGT GAG ATC AAA CCA GGA 4781
 Gln Glu Leu Tyr Thr Ile Glu Leu Pro Asn Asn Val Ser Glu Ile Lys Pro Gly
 CAA GAT GTA ACA GTC GTC ACA AAC AAT GGC AAA TCT TTC ACA TGT ACA CTC CGA 4835
 Gln Asp Val Thr Val Val Thr Asn Asn Gly Lys Ser Phe Thr Cys Thr Leu Arg
 TTT GAC ACA GAG gtaagttatctctagatgatataatttcgtcttttcgtcttttgattttttttt 4902
 Phe Asp Thr Glu
 agctctaaaaacaaaatgagatgagataagtaagaggaatttggaaatcgacagGTG GAG TTG GCT TAT 4969
 Val Glu Leu Ala Tyr
 TTC GAT CAC GGA GGG ATT TTG CAA TAC GTT ATC AGG AAC TTG ATC AAA CAA taa 5023
 Phe Asp His Gly Gly Ile Leu Gln Tyr Val Ile Arg Asn Leu Ile Lys Gln
 atctggtaacaccagagacttgagttatatatcccaagggttggttgcaataaaaatgtttgaggaggagg 5095
 acgagaatgactttaattttaacttttgccttttgccttctctcgtctcttgcttcgtgctagtcattggaa 5167
 acattatcactgtcgagtcatttttttctttcaaataaacatgagagctcttttttggttggttggttca 5239
 taagcattgtcaatgctgcattgatagaagtatactctaccattagaaaataaacacaaatacacaaagtaa 5311
 tctaaagagctagaggatgaaaattatctgtgaagggtgtacaaaaacattaaaaaaatctgagatgatcc 5383
 aaaggattgattatcacattcgaagtgatggttcagagattacctctggaatggtgtgtagagctgatgaac 5455
 tagagagcgacatga 5470

FIGURE 2(i)

6/36

ARABACO	ASPTPLINFLSSENRSLYFASSLDLYLSSMASENPFRSILKALEKPDGGE	50
MELONACO	HE-----	2
IREB_HUMAN	-----	0
IREB_MOUSE	MK-----NPFHAHLAEPLDAAQPGK	19
IREB_RABIT	MS-----NPFAYLAEPDPAQPGK	19
ACON_ECOLI	SST-----LREASK-----DTLQ-----AKDKTYHYYSPLAASKLGD	33
ACON_PIG	MAPYSL--VTRLQKAL---GVRQYHVASVLCQRA-KVAMSHFEPHEYIR	44
ACON_YEAST	MLSARSA--IKRPI-----VRGLATVSNLTRDS-KVNQNLLDHSFIN	40
ARABACO	FGNYYSLPALNDPRIDK-LPYSIRILLESAIRNCDEFQVKSXDVEKILDW	99
MELONACO	-----	2
IREB_HUMAN	-----CP---RKTRTQNLPPFW	13
IREB_MOUSE	--RFFNLNKLEDSRYGR-LPFSIRVLLLEAAVRNCDEFLVKKNDIENILNW	66
IREB_RABIT	--KFFNLNKLDYSRYGR-LPFSIRVLLLEAAVRNCDFLKKEDIDENILNW	66
ACON_ECOLI	IT-----R-LPKSLKVLLLENLLRWQDGNVTEEDIHALAGW	68
ACON_PIG	YDLLEKNIDIVRKRLNRPLTLSEKIVYCHLDD-----PANQETIERGKTY	88
ACON_YEAST	YKQNVETLDIVRKRLNRPFYAEKILYGHLLD-----PHGQDIQRGVSY	84
ARABACO	↓ ↓	
MELONACO	-ENTSPKQVEIPFKPARVLLQDFTGVPVAVVDLACMRDAMNLLGGDSNKIN	148
IREB_HUMAN	LSNKLKYNIEVPFKPARVILQDFTGVPVAVVDFAAMRDAVKKLGGDPEKIN	2
IREB_MOUSE	-NVMQHKNIIEVPFKPARVILQDFTGVPVAVVDFAAMRDAVKKLGGNPEKIN	115
IREB_RABIT	-NVTQHMNIIEVPFKPARVILQDFTGVPVAVVDFAAMRDAVKKLGGDPEKIN	115
ACON_ECOLI	LKNA-HADREIAYRPARVLMQDFTGVPVAVVDLAAMREAVKRLGGDTAKVN	117
ACON_PIG	LR-----LRPDRVAMQDATAQAMMLQF--ISSGLPKVAV-----	120
ACON_YEAST	LK-----LRPDRVACQDATAQMAILQF--MSAGLPQVAK-----	116
ARABACO	↓ ↓	
MELONACO	PLVPVDLVIDYSVQVDVARSENNAVQANMELEFQRNKEF AFLKWSNAFH	198
IREB_HUMAN	-----AKTENAVQANMELEFKRNRERFGLKWSGSAFH	35
IREB_MOUSE	PVCPADLVIDHSIQVDFNRRADSLQKNQDLEFERNRERFEFLKWSQAFH	113
IREB_RABIT	PVCPADLVIDHSIQVDFNRRADSLQKNQDLEFERNRERFEFLKWSQAFH	165
ACON_ECOLI	PLSPVDLVIDHSVTVDRFGDDEAFENVRLEMERNHERYVFLKWSQAFS	167
ACON_PIG	---PSTIHCDHLIEAQLGGEKD-LRRAKDI---NQEVYNFLATAGAKY-	161
ACON_YEAST	---PVTVHCDDLIAQAVGGEKD-LKRAIDL---NKEVYDFLASATAKY-	157
ARABACO	↓ ↓ ↓ ↓	
MELONACO	NMLVVP PGSGIVHQVNLEYLARVVFNT--NG--LLYPDSV-VGTDSTHTM	243
IREB_HUMAN	NMLVVP PGSGIVHQVNLEYLGRVVFNT--NG--LLYPDSV-VGTDSTHTM	80
IREB_MOUSE	NMRIIP PGSGIIHQVNLEYLARVVFQD--DG--YYYPDSL-VGTDSTHTM	158
IREB_RABIT	NMRIIP PGSGIIHQVNLEYLARVVFQD--DG--YYYPDSL-VGTDSTHTM	210
ACON_ECOLI	RFVVP PGSGIICHQVNLEYLGKAVWSELQDGEWLAYPELTL-VGTDSTHTM	216
ACON_PIG	GVGFWR PGSGIIHQIVLEN-----YAYPGVLLIGTDSHTPN	197
ACON_YEAST	NMGFWK PGSGIIHQIVLEN-----YAFPGALIIGTDSHTPN	193
ARABACO	↓ ↓ ↓ ↓	
MELONACO	IDGLGVAGWGVGGIEAERPM LGQ PMSMVLPGVVGFKLTGKLRDGMTATDL	293
IREB_HUMAN	IDGLGVAGWGVGGIEAEAAM LGQ PMSMVLPGVVGFKLVGKL RGV TATDL	130
IREB_MOUSE	IDGLGILGWGVGGIEAEAAM LGQ PISMVLPQVIGYRLMGKPHPLVTSTDI	208
IREB_RABIT	IDGLGVLGWGVGGIEAEAAM LGQ PISMVLPQVIGYRLMGKPHPLVTSTDI	260
ACON_ECOLI	INGLGVLGWGVGGIEAEAAM LGQ PVSMLIPDVVGFKLTGKLRGEGITATDL	266
ACON_PIG	GGGLGGICIGVGGADAVDVMAGIPWELKCPKVIKLTGSLSGWTSPPKDV	247
ACON_YEAST	AGGLGQLAIGVGGADAVDVMAGIPWELKAPKILGKLTGKMGWTSPPKDI	243
ARABACO	↓ ↓	
MELONACO	VLTVTQMLRKHG VVGK FVEFHGEGMREL SLADRATIANMSPEYGATMGFF	343
IREB_HUMAN	VLTVTQMLRKHG VVGK FVEFHGEGMREL SLADRATIANMSPEYGATMGFF	180
IREB_MOUSE	VLTVTQMLRKHG VVGK FVEFFGPGVAQLSIADRATIANMCPYGGATAFF	258
IREB_RABIT	VLTVTQMLRKHG VVGK FVEFFGPGVAQLSIADRATIANMCPYGGATAFF	310
ACON_ECOLI	VLTVTQMLRKHG VVGK FVEFFGPGVAQLSIADRATIANMCPYGGATAFF	310
ACON_PIG	ILKVAGILTVKGGTGAI VYHGPVDSISCTGMATICNMGAIEIGATTSVF	297
ACON_YEAST	ILKLAGITTVKGGTGKIVYFGDGVDTFSATGMATICNMGAIEIGATTSVF	293

FIGURE 2 (ii)

ARABACO	PVDHVTLLQYLRLTGRSDDTVSMIEAYLRANKMFVDYSEPESEKTVYSSCLE	393
MELONACO	PVDHVTLLQYLKLTGRKDETISMIESYLLANKMFVDYSEPOVERVYSSHIE	230
IRES_HUMAN	PVDEVSITYLVQTRDEEKLKYIKYLQAVGMFRDFNDPSQDPDFTQVVE	308
IRES_MOUSE	PVDEVSIAAYLLQTRDEEDKVKHIQKYLQAVGMFRDFNDTSQDPDFTQVVE	360
IRES_RABIT	PVDEVSIIKYLQTRDESKVKQIRKYLQAVGMFRDYSQDPDFTQVVE	360
ACON_ECOLI	PIDAVTLDYMLSGRSEDQVELVEKYAKAQGMWRN---PGDEPIFTSTLE	363
ACON_PIG	PYNHRMKKYLSTKGRADIA-NLADEF-X-----DHLVDPGCHYDQVIE	339
ACON_YEAST	PFNKSMIEYLEATGRCKIA-DFAKLYHK-----DLLSADKDAEYDEVVE	336
*		
ARABACO	LNLSDVEPCISGPKRPHDRVPLKEMKADWHSCLDNRVGFKGFAVPKEAQS	443
MELONACO	LNLSDVEPCISGPKRPHDRVPLKEMKADWHACLDNRVGFKGFAVPKEAQS	280
IRES_HUMAN	LDLKTVPVCCSGPKRPQDKVAVSDMKDFESCLGAKQGFQVAPDRHN	358
IRES_MOUSE	LDLKTVPVCCSGPKRPQDKVAVSEMCKDFESCLGAKQGFQVAPDRHN	410
IRES_RABIT	LDLKTVPVCCSGPKRPQDKVAVSDMKDFESCLGAKQGFQVAPDRHN	410
ACON_ECOLI	LDMDVVEASLAGPKRPQDRVALPDVPAFAASNELEVN-----ATHKDR	407
ACON_PIG	INLSELKPHINGPFTPLAHPVAEVGS-----	366
ACON_YEAST	IDLNTLEPYINGPFTPLATPVSKMKE-----	363
*		
ARABACO	KAVEFNFNNGTTAQLRHGDVVIAAITSCTNTSNPSVMLGAALVAKKACDLG	493
MELONACO	KVAEFNFHGSQAQLRHGDVVIAAITSCTNTSNPSVMLGAALVAKKACELG	330
IRES_HUMAN	DHKTFIYDNTEFTLAHGSVVIAAITSCTNTSNPSVMLGAGLLAKKAVDAG	408
IRES_MOUSE	DRKTFILYNSSEFTLAHGSVVIAAITSCTNTSNPSVMLGAGLLAKKAVEAG	460
IRES_RABIT	DHKTFIYDNSEFTLSHGSVVIAAITSCTNTSNPSVMLGAGLLAKKAVDAG	460
ACON_ECOLI	QPVYVMNGHQYQLPDGAVVIAAITSCTNTSNPSVMAAGLLAKKAVTLG	457
ACON_PIG	-----VAKEGKWL---DIRVGLIGSCTNSSYED-MGRSAAVAKQALAHG	407
ACON_YEAST	-----VAVANNWPL---DVRVGLIGSCTNSSYED-MSRSASIVKDAAHG	404
*		
ARABACO	LEVKPIWIKTSLAPGSGVVTYKYLAKSGLQKYLNLQGFSTVGYGCTTCIGNS	543
MELONACO	LEVKPIWIKTSLAPGSGVVTYKYLAKSGLQKYLNLQGFSTVGYGCTTCIGNS	380
IRES_HUMAN	LNVMPYIKTSLSPGSGVVTYLLQESGVMPYLSQLGFDVVGYGCTTCIGNS	458
IRES_MOUSE	LSVKPYIKTSLSPGSGVVTYLLRESGVMPYLSQLGFDVVGYGCTTCIGNS	510
IRES_RABIT	LNVKPYIKTSLSPGSGVVTYLLRESGVMPYLSQLGFDVVGYGCTTCIGNS	510
ACON_ECOLI	LKRQPVWKASLAPGSKVSDYLAKAKLTPYLDELGFNLVGYGCTTCIGNS	507
ACON_PIG	LKCKSQF---TITPGSEQIRATIERDGYAQVLRDVGSIIVLANACGPGCIGQW	455
ACON_YEAST	LKSKTIF---TVTPGSEQIRATIERDQLETKEFGGIIVLANACGPGCIGQW	452
*		
ARABACO	GDIHEAVASAIVDNDLVASAVLSGNRNFEGRVHPLTRAN-YLASPPLVVA	592
MELONACO	GDIDESVASAITGNDIVAAAVLSGNRNFEGRVHPLTRAN-YLASPPLVVA	429
IRES_HUMAN	GPLPEPVVEAITQGDILVAVGVIIIGTGILKAEIYPNTRAN-YLASPPLVIA	507
IRES_MOUSE	GPLPEPVVEAITQGDILVAVGVLSGNRNFEGRVHPLTRAN-YLASPPLVIA	559
IRES_RABIT	GPLPEPVVEAITQGDILVAVGVLSGNRNFEGRVHPLTRAN-YLASPPLVIA	559
ACON_ECOLI	GPLPDPITETAIKKGDLTVGAVLSGNRNFEGRVHPLVKTN-WLASPPLVVA	556
ACON_PIG	DR-----KDICKGEKNTI-VTSYNRNFTGRNDANPETHAFVTSPEIVTA	498
ACON_YEAST	DR-----RDIKKGDKNTI-VSSYNRNFTSRNDGNPQTHAFVASPEIVTA	495
*		
ARABACO	YALAGTVDDIDFETQPIGTGKDGKQIFFRDIWPSNKEVAEVVQSSVLPDMF	642
MELONACO	YALAGTVDDIDFESEPIGVGKDGKQVFFRDIWPTSZEVAVVNSNVLPMF	479
IRES_HUMAN	YAIAGTIRIDFEKEPLGVNAKGQOVFLKDIWPTREDEIQAVEROQVPIPGMF	557
IRES_MOUSE	YAIAGTVRIDFEKEPLGVNAKGQOVFLKDIWPTREDEIQAVEROQVPIPGMF	609
IRES_RABIT	YAIAGTIRIDFEKEPLGTNAKGQOVFLRDIWPTREEIQAVEROQVPIPGMF	609
ACON_ECOLI	YALAGNMNINLASEPIGHDRKGDVYLLKDIWPSAQEIARAVEQ-VSTEMF	605
ACON_PIG	LAIAGTLKFNPETDFL-TGKDGKGFLEA--PDADLPRAEFDPGQDTYQ	545
ACON_YEAST	FAIAGDLRFNPLTDKL-KDKDGNEFMKLP--PHGRWFASKEVMMLVRTL	542
*		
ARABACO	KATYEAITKGNMWNQLSVASGTLIYWDPKSTYIHEPPYFKGMTMSPPGP	692
MELONACO	RATYQAITTEGNATWNLLSVPEGTLYSWDPTSTYIHEPPYFKDMSMSPPGP	529
IRES_HUMAN	KEVYQKIETVNESWNALATPSDKLFFWNSKSTYIKSPFFENLTLDLOPP	607
IRES_MOUSE	KEVYQKIETVKNKSWNALAAPSEKLYAWNPKSTYIKSPFFESLTLDLOPP	659
IRES_RABIT	TEVYQKIETVNASWNALAAPSKLYLWNPKSTYIKSPFFENLTLDLOPP	659
ACON_ECOLI	RKEYAEVFEGTAEWKGINVTRSDTYGWQEDSTYIRLSPPFFDEMOPATPAV	655
ACON_PIG	HPPKDSGQ-RVD---VSPTSRLQLLEP-----FDKWD-----G	576
ACON_YEAST	KLHLQTVATVEVK---VSPTSRLQLLKP-----FKPWD-----G	574

FIGURE 2 (iii)

ARABACO	HGVKDAYCLLNFGDSITTDHISPAGSIHKDSPAAYLMERGVDRRDFNSY	742
MELONACO	HGVKNAYCLLNFGDSITTDHISPAGSIHKDSPAAYLLERGVDRRDFNSY	579
IREB_HUMAN	KSIVDAYVLLNLGDSVTTDHI SPAGNIARNSPAARYLTNRGLTPREFNSY	657
IREB_MOUSE	KSIVDAYVLLNLGDSVTTDHI SPAGNIARNSPAARYLTNRGLTPREFNSY	709
IREB_RABIT	KSIVDAYVLLNLGDSVTTDHI SPAGNIARNSPAARYLTNRGLTPREFNSY	709
ACON_ECOLI	EDIHGARILAMLGDSVTTDHI SPAGSIKPDSPAGRYLQGRGVERKDFNSY	705
ACON_PIG	KDLEDLQILIKVKGKCTTDHISAAGP-----WLKFRGHLDNISNNL	617
ACON_YEAST	KDAKDMFILIKAVGKTTTDHISMAGP-----WLKYRGHLENISNNY	615
..
ARABACO	GVAVVMRLW-REHFANIRIVNKLKGEVGPKTWHIPTGEXLSVFDAAAMK	791
MELONACO	GVAVVMRLW-HVHFANIRIVNKLKGEVGPKTIHIPREKLSVFDAAAMR	628
IREB_HUMAN	GSPEVMTPSWGHGHLPTLRLLNRFNLKQ-APQTIHLPSGETLDFDAAER	706
IREB_MOUSE	GSRRGNDAIMARGTFANIRLLNRFNLKQ-APQTIHLPSGETLDFDAAER	758
IREB_RABIT	GSRRGNDAIMARGTFANIRLLNRFNLKQ-APQTIHLPSGETLDFDAAER	758
ACON_ECOLI	GSRRGNHEVMMRGTFANIRIRNEMVPGVEGGMTRHLPSDVVSVYDAAMR	755
ACON_PIG	-----LIGAINIENR-----KANSVRNAVTOEFGVPPTARY	649
ACON_YEAST	-----MIGAINAENK-----KANCYKNVYTGZKGVPTARD	647
..
ARABACO	YRNEGRDTIILAGAEYGGSGSRDWAAGPMLLGKAVISKSFERIHRSNL	841
MELONACO	YKSEGQDTIILAGAEYGGSGSRDWAAGPMLLGKAVIAKTFERIHRSNL	678
IREB_HUMAN	YQQAGLPLIVLAGKEYGAGSSRDWAAGPFLGKAVLAESYERIHRSNL	756
IREB_MOUSE	YQQAGLPLIVLAGKEYGAGSSRDWAAGPFLGKAVLAESYERIHRSNL	808
IREB_RABIT	YQQEGHPLIVLAGKEYGAGSSRDWAAGPFLGKAVLAESYERIHRSNL	808
ACON_ECOLI	YKQEQTPLAIVLAGKEYGAGSSRDWAAGPRLGIRVVIASFERIHRSNL	805
ACON_PIG	YKQHGIRWVIGDENYGGSSREHRALEPRHLGGRAITKSFARIHETNL	699
ACON_YEAST	YRDQGIKWVIGDENFEGGSSREHAALPRFLGGFAITKSFARIHETNL	697
..
ARABACO	VGMGIIPLCFKAGZDAETLGLTGQELVTIELPMNVSEIKFGQDVTVTNN	891
MELONACO	VGMGIIPLCFKAGZDADSLGLTGHERFTIDLPSNVGEIRPGQDVAVVTD	728
IREB_HUMAN	VGMGVIPLEVLPGENADALGLTGQERYTIIPIE---NLKPFQMKVQVCLDT	803
IREB_MOUSE	VGMGVIPLEVLPGETADSLGLTGRERYTINPIE---DLKPRMTVQIKLDT	855
IREB_RABIT	VGMGVIPLEVLPGENADSLGLTGRERYTIIPIE---NLTPRMHVQVCLDT	855
ACON_ECOLI	IGMGILPLEFPQGVTRKTLGLTGEEKIDI---GDLQNLQPGATVPVTLTR	852
ACON_PIG	KKQGLLPLTFADPADYDKI---HPVDKLTIQ---GLKDFAPGKPLACITKH	744
ACON_YEAST	KKQGLLPLNFKNPADYDKI---NPDDRIDL---GLAELAPGKPVTVRVHP	742
..
ARABACO	GKSFTCTL----RFDTEVELAYFDHGGILQYVIRNLI--KQ	926
MELONACO	GKSPFSCIL----RFDTEVELAYFDHGGILQYVIRNLIHSHK	765
IREB_HUMAN	GKTFQAVM----RFDTDVELTYFLNGGILNVMIRKMA---K	837
IREB_MOUSE	GKTFQAVM----RFDTDVELTYFHNGGILNVMIRKMA---Q	889
IREB_RABIT	GKTFQAVI----RFDTDVELTYFLNGGILNVMIRKMA---K	889
ACON_ECOLI	ADGSQEVVPCRCRIDTATELTYYQNDGILEYVIRNML---K	890
ACON_PIG	PNGTOETILLNHTFN--ETQLEWFRAGSALN---RMKELQOK	781
ACON_YEAST	KNGKXPWDAVLTHTFN--DEQLEWFKYGSALN---KKADEKX	779
..

FIGURE 3 (i)

																	a	-1
AAC	TTT	CTC	TCT	TCT	GAA	AAC	AGA	TCT	CTC	TAC	TTT	GCT	TCT	TCA	CTC	GAC	TTG	54
Asn	Phe	Leu	Ser	Ser	Glu	Asn	Arg	Ser	Leu	Tyr	Phe	Ala	Ser	Ser	Leu	Asp	Leu	17
TAT	CTA	TCA	TCC	ATG	GCT	TCC	GAG	AAT	CCT	TTC	CGA	AGC	ATA	TTG	AAG	GCG	TTA	108
Tyr	Leu	Ser	Ser	Met	Ala	Ser	Glu	Asn	Pro	Phe	Arg	Ser	Ile	Leu	Lys	Ala	Leu	35
GAG	AAG	CCT	GAT	GGT	GGT	GAA	TTC	GGT	AAC	TAC	TAC	AGC	TTA	CCT	GCT	TTG	AAC	162
Glu	Lys	Pro	Asp	Gly	Gly	Glu	Phe	Gly	Asn	Tyr	Tyr	Ser	Leu	Pro	Ala	Leu	Asn	53
GAT	CCC	AGG	ATC	GAT	AAA	CTA	CCT	TAT	TCC	ATT	AGG	ATA	CTT	CTT	GAA	TCG	GCC	216
Asp	Pro	Arg	Ile	Asp	Lys	Leu	Pro	Tyr	Ser	Ile	Arg	Ile	Leu	Leu	Glu	Ser	Ala	71
ATA	CGT	AAC	TGT	GAT	GAG	TTC	CAA	GTT	AAG	AGC	AAA	GAT	GTT	GAG	AAG	ATT	CTT	270
Ile	Arg	Asn	Cys	Asp	Glu	Phe	Gln	Val	Lys	Ser	Lys	Asp	Val	Glu	Lys	Ile	Leu	89
GAT	TGG	GAG	AAT	ACT	TCT	CCC	AAG	CAG	GTT	GAG	ATT	CCG	TTC	AAG	CCT	GCT	CGG	324
Asp	Trp	Glu	Asn	Thr	Ser	Pro	Lys	Gln	Val	Glu	Ile	Pro	Phe	Lys	Pro	Ala	Arg	107
GTT	CTT	CTT	CAG	GAC	TTT	ACT	GGT	GTT	CCT	GCT	GTT	GTT	GAT	CTT	GCT	TGC	ATG	378
Val	Leu	Leu	Gln	Asp	Phe	Thr	Gly	Val	Pro	Ala	Val	Val	Asp	Leu	Ala	Cys	Met	125
AGA	GAT	GCC	ATG	AAT	AAT	CTC	GGT	GGT	GAT	TCT	AAT	AAA	ATT	AAT	CCG	CTG	GTC	432
Arg	Asp	Ala	Met	Asn	Asn	Leu	Gly	Gly	Asp	Ser	Asn	Lys	Ile	Asn	Pro	Leu	Val	143
CCT	GTA	GAT	CTT	GTC	ATT	GAC	TAC	TCC	GTT	CAG	GTG	GAT	GTG	GCG	AGA	TCA	GAG	486
Pro	Val	Asp	Leu	Val	Ile	Asp	Tyr	Ser	Val	Gln	Val	Asp	Val	Ala	Arg	Ser	Glu	161
AAC	GCA	GTG	CAG	GCA	AAC	ATG	GAG	CTT	GAG	TTC	CAG	CGT	AAC	AAG	GAA	AGA	TTT	540
Asn	Ala	Val	Gln	Ala	Asn	Met	Glu	Leu	Glu	Phe	Gln	Arg	Asn	Lys	Glu	Arg	Phe	179
GCT	TTT	CTT	AAG	TGG	GGA	TCC	AAC	GCC	TTT	CAC	AAC	ATG	CTT	GTC	GTA	CCT	CCT	594
Ala	Phe	Leu	Lys	Trp	Gly	Ser	Asn	Ala	Phe	His	Asn	Met	Leu	Val	Val	Pro	Pro	197
GGA	TCT	GGA	ATA	GTT	CAT	CAA	GTC	AAC	CTA	GAA	TAC	CTT	GCC	AGA	GTT	GTT	TTC	648
Gly	Ser	Gly	Ile	Val	His	Gln	Val	Asn	Leu	Glu	Tyr	Leu	Ala	Arg	Val	Val	Phe	215
AAC	ACA	AAT	GGA	CTT	CTT	TAC	CCA	GAC	AGT	GTT	GTT	GGC	ACA	GAC	TCT	CAC	ACC	702
Asn	Thr	Asn	Gly	Leu	Leu	Tyr	Pro	Asp	Ser	Val	Val	Gly	Thr	Asp	Ser	His	Thr	233
ACT	ATG	ATT	GAT	GGA	CTG	GGT	GTT	GCT	GGA	TGG	GGA	GTT	GGC	GGT	ATA	GAA	GCG	756
Thr	Met	Ile	Asp	Gly	Leu	Gly	Val	Ala	Gly	Trp	Gly	Val	Gly	Gly	Ile	Glu	Ala	251
GAA	CGA	CCG	ATG	CTT	GGT	CAG	CCA	ATG	AGC	ATG	GTC	CTA	CCC	GGT	GTT	GTG	GGT	810
Glu	Arg	Pro	Met	Leu	Gly	Gln	Pro	Met	Ser	Met	Val	Leu	Pro	Gly	Val	Val	Gly	269
TTC	AAG	CTA	ACG	GGA	AAG	TTA	AGA	GAT	GGA	ATG	ACA	GCT	ACT	GAT	TTG	GTC	TTA	864
Phe	Lys	Leu	Thr	Gly	Lys	Leu	Arg	Asp	Gly	Met	Thr	Ala	Thr	Asp	Leu	Val	Leu	287
ACA	GTG	ACT	CAG	ATG	TTG	AGG	AAA	CAT	GGA	GTA	GTT	GGA	AAG	TTT	GTT	GAA	TTC	918
Thr	Val	Thr	Gln	Met	Leu	Arg	Lys	His	Gly	Val	Val	Gly	Lys	Phe	Val	Glu	Phe	305
CAC	GGG	GAA	GGG	ATG	AGA	GAA	TTG	TCT	TTA	GCT	GAC	CGT	GCT	ACA	ATT	GCC	AAT	972
His	Gly	Glu	Gly	Met	Arg	Glu	Leu	Ser	Leu	Ala	Asp	Arg	Ala	Thr	Ile	Ala	Asn	323
ATG	TCT	CCT	GAG	TAC	GGT	GCG	ACC	ATG	GGA	TTC	TTC	CCA	GTC	GAT	CAT	GTC	ACT	1026
Met	Ser	Pro	Glu	Tyr	Gly	Ala	Thr	Met	Gly	Phe	Phe	Pro	Val	Asp	His	Val	Thr	341
TTG	CAG	TAT	CTA	AGG	TTG	ACA	GGC	AGG	AGC	GAT	GAC	ACT	GTC	TCC	ATG	ATA	GAG	1080
Leu	Gln	Tyr	Leu	Arg	Leu	Thr	Gly	Arg	Ser	Asp	Asp	Thr	Val	Ser	Met	Ile	Glu	359

FIGURE 3 (ii)

GCG	TAT	TTA	CGA	GCA	AAC	AAG	ATG	TTT	GTG	GAT	TAC	AGT	GAG	CCG	GAG	AGT	AAG	1134
Ala	Tyr	Leu	Arg	Ala	Asn	Lys	Met	Phe	Val	Asp	Tyr	Ser	Glu	Pro	Glu	Ser	Lys	377
ACA	GTT	TAT	TCC	TCA	TGT	CTG	GAA	TTG	AAT	CTC	GAG	GAT	GTG	GAA	CCT	TGT	GTT	1188
Thr	Val	Tyr	Ser	Ser	Cys	Leu	Glu	Leu	Asn	Leu	Glu	Asp	Val	Glu	Pro	Cys	Val	395
TCT	GGT	CCC	AAG	AGG	CCT	CAT	GAT	CGT	GTT	CCT	TTG	AAG	GAA	ATG	AAA	GCG	GAC	1242
Ser	Gly	Pro	Lys	Arg	Pro	His	Asp	Arg	Val	Pro	Leu	Lys	Glu	Met	Lys	Ala	Asp	413
TGG	CAT	TCT	TGC	TTG	GAC	AAT	AGA	GTA	GGA	TTC	AAG	GGT	TTC	GCT	GTA	CCT	AAA	1296
Trp	His	Ser	Cys	Leu	Asp	Asn	Arg	Val	Gly	Phe	Lys	Gly	Phe	Ala	Val	Pro	Lys	431
GAA	GCA	CAG	AGT	AAG	GCT	GTA	GAG	TTC	AAT	TTT	AAC	GGG	ACC	ACA	GCA	CAG	CTT	1350
Glu	Ala	Gln	Ser	Lys	Ala	Val	Glu	Phe	Asn	Phe	Asn	Gly	Thr	Thr	Ala	Gln	Leu	449
AGA	CAT	GGA	GAT	GTT	GTT	ATA	GCA	GCA	ATC	ACC	AGT	TGC	ACA	AAT	ACT	TCA	AAC	1404
Arg	His	Gly	Asp	Val	Val	Ile	Ala	Ala	Ile	Thr	Ser	Cys	Thr	Asn	Thr	Ser	Asn	467
CCT	AGT	GTA	ATG	CTT	GGC	GCT	GCC	TTA	GTT	GCA	AAA	AAG	GCC	TGC	GAC	CTA	GGA	1458
Pro	Ser	Val	Met	Leu	Gly	Ala	Ala	Leu	Val	Ala	Lys	Lys	Ala	Cys	Asp	Leu	Gly	485
CTG	GAG	GTT	AAG	CCA	TGG	ATC	AAA	ACT	AGT	CTT	GCT	CCA	GGC	TCT	GGA	GTT	GTA	1512
Leu	Glu	Val	Lys	Pro	Trp	Ile	Lys	Thr	Ser	Leu	Ala	Pro	Gly	Ser	Gly	Val	Val	503
ACA	AAG	TAC	TTG	GCA	AAG	AGT	GGC	TTG	CAG	AAG	TAC	TTG	AAT	CAG	CTC	GGC	TTC	1566
Thr	Lys	Tyr	Leu	Ala	Lys	Ser	Gly	Leu	Gln	Lys	Tyr	Leu	Asn	Gln	Leu	Gly	Phe	521
AGT	ATC	GTT	GGT	TAT	GGG	TGC	ACC	ACA	TGC	ATT	GGA	AAC	TCG	GGG	GAT	ATC	CAT	1620
Ser	Ile	Val	Gly	Tyr	Gly	Cys	Thr	Thr	Cys	Ile	Gly	Asn	Ser	Gly	Asp	Ile	His	539
GAA	GCT	GTG	GCT	TCA	GCA	ATA	GTT	GAT	AAT	GAC	TTG	GTG	GCA	TCC	GCT	GTG	TTG	1674
Glu	Ala	Val	Ala	Ser	Ala	Ile	Val	Asp	Asn	Asp	Leu	Val	Ala	Ser	Ala	Val	Leu	557
TCT	GGG	AAC	AGA	AAT	TTT	GAG	GGA	CGT	GTT	CAC	CCG	TTA	ACA	AGA	GCT	AAC	TAT	1728
Ser	Gly	Asn	Arg	Asn	Phe	Glu	Gly	Arg	Val	His	Pro	Leu	Thr	Arg	Ala	Asn	Tyr	575
CTA	GCT	TCC	CCA	CCG	CTT	GTT	GTA	GCC	TAT	GCT	CTG	GCT	GGG	ACT	GTT	GAC	ATT	1782
Leu	Ala	Ser	Pro	Pro	Leu	Val	Val	Ala	Tyr	Ala	Leu	Ala	Gly	Thr	Val	Asp	Ile	593
GAT	TTT	GAG	ACA	CAG	CCC	ATT	GGA	ACT	GGG	AAA	GAT	GGA	AAA	CAG	ATA	TTT	TTC	1836
Asp	Phe	Glu	Thr	Gln	Pro	Ile	Gly	Thr	Gly	Lys	Asp	Gly	Lys	Gln	Ile	Phe	Phe	611
AGG	GAC	ATT	TGG	CCC	TCT	AAC	AAA	GAA	GTT	GCT	GAG	GTT	GTT	CAA	TCT	AGT	GTC	1890
Arg	Asp	Ile	Trp	Pro	Ser	Asn	Lys	Glu	Val	Ala	Glu	Val	Val	Gln	Ser	Ser	Val	629
CTT	CCT	GAT	ATG	TTC	AAA	GCT	ACA	TAT	GAA	GCA	ATC	ACC	AAA	GGA	AAT	TCC	ATG	1944
Leu	Pro	Asp	Met	Phe	Lys	Ala	Thr	Tyr	Glu	Ala	Ile	Thr	Lys	Gly	Asn	Ser	Met	647
TGG	AAT	CAG	TTA	TCT	GTG	GCG	TCA	GGT	ACT	CTC	TAT	GAG	TGG	GAC	CCG	AAA	TCA	1998
Trp	Asn	Gln	Leu	Ser	Val	Ala	Ser	Gly	Thr	Leu	Tyr	Glu	Trp	Asp	Pro	Lys	Ser	665
ACT	TAC	ATT	CAC	GAG	CCG	CCT	TAT	TTC	AAG	GGC	ATG	ACC	ATG	TCT	CCA	CCC	GGT	2052
Thr	Tyr	Ile	His	Glu	Pro	Pro	Tyr	Phe	Lys	Gly	Met	Thr	Met	Ser	Pro	Pro	Gly	683
CCA	CAT	GGT	GTG	AAA	GAC	GCA	TAC	TGT	TTA	CTC	AAT	TTT	GGA	GAC	AGT	ATT	ACC	2106
Pro	His	Gly	Val	Lys	Asp	Ala	Tyr	Cys	Leu	Leu	Asn	Phe	Gly	Asp	Ser	Ile	Thr	701
ACT	GAT	CAC	ATC	TCA	CCA	GCT	GGT	AGC	ATC	CAC	AAG	GAC	AGT	CCT	GCG	GCT	AAG	2160
Thr	Asp	His	Ile	Ser	Pro	Ala	Gly	Ser	Ile	His	Lys	Asp	Ser	Pro	Ala	Ala	Lys	719
TAC	TTG	ATG	GAA	CGA	GGT	GTG	GAT	AGA	AGA	GAC	TTC	AAC	TCA	TAC	GGA	GTC	GCC	2214
Tyr	Leu	Met	Glu	Arg	Gly	Val	Asp	Arg	Arg	Asp	Phe	Asn	Ser	Tyr	Gly	Val	Ala	737

11/36

FIGURE 3 (iii)

[illegible]

12/36

FIGURE 4 (i)

																	gg	-1
CAC	GAG	GCA	AAA	ACA	GAA	AAT	GCA	GTA	CAG	GCA	AAT	ATG	GAA	CTT	GAA	TTT	AAA	54
His	Glu	Ala	Lys	Thr	Glu	Asn	Ala	Val	Gln	Ala	Asn	Met	Glu	Leu	Glu	Phe	Lys	16
CGA	AAC	AGG	GAA	AGA	TTT	GGC	TTT	CTG	AAA	TGG	GGA	TCG	AGT	GCT	TTC	CAC	AAC	108
Arg	Asn	Arg	Glu	Arg	Phe	Gly	Phe	Leu	Lys	Trp	Gly	Ser	Ser	Ala	Phe	His	Asn	34
ATG	CTT	GTC	GTA	CCT	CCA	GGA	TCT	GGT	ATA	GTT	CAC	CAG	GTT	AAT	CTG	GAA	TAC	162
Met	Leu	Val	Val	Pro	Pro	Gly	Ser	Gly	Ile	Val	His	Gln	Val	Asn	Leu	Glu	Tyr	52
CTT	GGC	CGA	GTG	GTA	TTC	AAC	ACA	AAC	GGT	TTG	CTT	TAC	CCT	GAT	AGT	GTT	GTG	216
Leu	Gly	Arg	Val	Val	Phe	Asn	Thr	Asn	Gly	Leu	Leu	Tyr	Pro	Asp	Ser	Val	Val	70
GGA	ACC	GAT	TCA	CAC	ACA	ACG	ATG	ATT	GAC	GGA	TTG	GGT	GTT	GCT	GGT	TGG	GGT	270
Gly	Thr	Asp	Ser	His	Thr	Thr	Met	Ile	Asp	Gly	Leu	Gly	Val	Ala	Gly	Trp	Gly	88
GTT	GGT	GGG	ATA	GAA	GCA	GAA	GCT	GCA	ATG	CTT	GGA	CAG	CCA	ATG	AGC	ATG	GTC	324
Val	Gly	Gly	Ile	Glu	Ala	Glu	Ala	Ala	Met	Leu	Gly	Gln	Pro	Met	Ser	Met	Val	106
TTG	CCT	GGC	GTT	GTT	GGT	TTT	AAA	TTA	GTT	GGA	AAG	CTG	AGA	AAT	GGT	GTA	ACA	378
Leu	Pro	Gly	Val	Val	Gly	Phe	Lys	Leu	Val	Gly	Lys	Leu	Arg	Asn	Gly	Val	Thr	124
GCT	ACA	GAC	TTG	GTT	TTG	ACT	GTG	ACT	CAA	ATG	CTC	AGG	AAG	CAT	GGT	GTT	GTT	432
Ala	Thr	Asp	Leu	Val	Leu	Thr	Val	Thr	Gln	Met	Leu	Arg	Lys	His	Gly	Val	Val	142
GGC	AAG	TTT	GTG	GAG	TTC	TAT	GGG	GAA	GGT	ATG	GGT	GAA	CTA	TCT	TTG	GCC	GAC	486
Gly	Lys	Phe	Val	Glu	Phe	Tyr	Gly	Glu	Gly	Met	Gly	Glu	Leu	Ser	Leu	Ala	Asp	160
CGT	GCC	ACA	ATT	GCC	AAC	ATG	TCC	CCT	GAA	TAT	GGT	GCA	ACA	ATG	GGA	TTC	TTT	540
Arg	Ala	Thr	Ile	Ala	Asn	Met	Ser	Pro	Glu	Tyr	Gly	Ala	Thr	Met	Gly	Phe	Phe	178
CCT	GTG	GAT	CAT	GTC	ACC	CTG	CAG	TAT	CTA	AAA	CTG	ACT	GGA	AGA	AAA	GAT	GAA	594
Pro	Val	Asp	His	Val	Thr	Leu	Gln	Tyr	Leu	Lys	Leu	Thr	Gly	Arg	Lys	Asp	Glu	196
ACA	ATT	TCT	ATG	ATA	GAA	TCT	TAC	CTG	CTG	GCT	AAT	AAG	ATG	TTT	GTA	GAC	TAC	648
Thr	Ile	Ser	Met	Ile	Glu	Ser	Tyr	Leu	Leu	Ala	Asn	Lys	Met	Phe	Val	Asp	Tyr	214
AGT	GAG	CCT	CAA	GTT	GAA	AGA	GTG	TAC	TCC	TCT	CAT	ATA	GAA	CTG	AAT	CTT	TCT	702
Ser	Glu	Pro	Gln	Val	Glu	Arg	Val	Tyr	Ser	Ser	His	Ile	Glu	Leu	Asn	Leu	Ser	232
GAT	GTC	GAA	CCA	TGC	ATA	TCA	GGT	CCA	AAA	AGA	CCA	CAT	GAT	CGA	GTT	CCC	TTG	756
Asp	Val	Glu	Pro	Cys	Ile	Ser	Gly	Pro	Lys	Arg	Pro	His	Asp	Arg	Val	Pro	Leu	250
AAA	GAG	ATG	AAA	GCT	GAT	TGG	CAT	GCA	TGC	CTT	GAC	AAC	AGA	GTT	GGA	TTC	AAG	810
Lys	Glu	Met	Lys	Ala	Asp	Trp	His	Ala	Cys	Leu	Asp	Asn	Arg	Val	Gly	Phe	Lys	268
GGT	TTT	GCC	ATA	CCA	AAG	GAA	GCT	CAA	GTA	AAG	GTT	GCA	GAG	TTC	AAT	TTT	CAC	864
Gly	Phe	Ala	Ile	Pro	Lys	Glu	Ala	Gln	Val	Lys	Val	Ala	Glu	Phe	Asn	Phe	His	286
GGA	TCC	CCA	GCT	CAG	CTT	AGG	CAT	GGG	GAT	GTT	GTA	ATT	GCA	GCT	ATC	ACT	AGC	918
Gly	Ser	Pro	Ala	Gln	Leu	Arg	His	Gly	Asp	Val	Val	Ile	Ala	Ala	Ile	Thr	Ser	304
TGC	ACA	AAT	ACC	TCT	TCC	AGC	GTT	ATG	CTT	GGA	GCT	GCT	TTG	GTG	GCT	AAA	AAG	972
Cys	Thr	Asn	Thr	Ser	Ser	Ser	Val	Met	Leu	Gly	Ala	Ala	Leu	Val	Ala	Lys	Lys	322
GCT	TGT	GAG	CTT	GGG	CTA	GAG	GTT	AAG	CCC	TGG	ATT	AAG	ACA	GTC	TTG	CTC	CAG	1026
Ala	Cys	Glu	Leu	Gly	Leu	Glu	Val	Lys	Pro	Trp	Ile	Lys	Thr	Val	Leu	Leu	Gln	340
GCT	CTG	GGG	GTG	GTG	ACG	AAA	TAT	TTG	GCG	AAG	AGT	GGC	TTG	CAA	AAG	TAT	CTT	1080
Ala	Leu	Gly	Val	Val	Thr	Lys	Tyr	Leu	Ala	Lys	Ser	Gly	Leu	Gln	Lys	Tyr	Leu	358
AAT	CAG	CTA	GGA	TTC	AAT	ATA	GTT	GGT	TAC	GGA	TGT	ACT	ACT	TGC	ATT	GGC	AAT	1134
Asn	Gln	Leu	Gly	Phe	Asn	Ile	Val	Gly	Tyr	Gly	Cys	Thr	Thr	Cys	Ile	Gly	Asn	376

FIGURE 4 (ii)

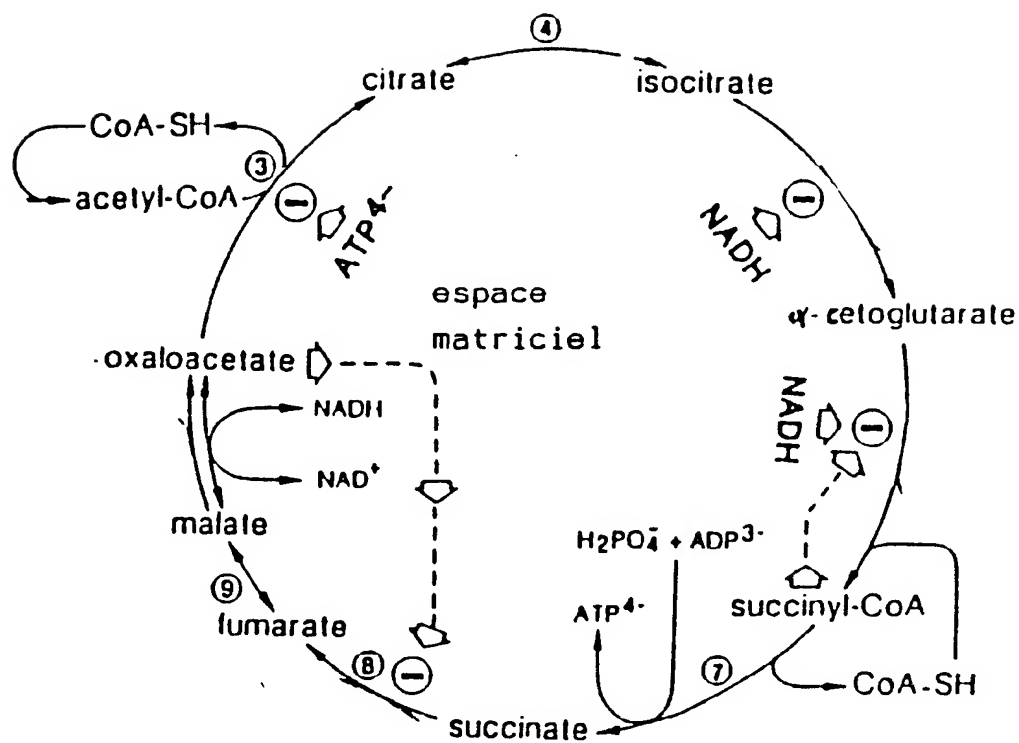
TCT	GGA	GAT	ATT	GAT	GAA	TCC	GTT	GCT	TCA	GCA	ATA	ACC	GGA	AAT	GAC	ATA	GTT	1188
Ser	Gly	Asp	Ile	Asp	Glu	Ser	Val	Ala	Ser	Ala	Ile	Thr	Gly	Asn	Asp	Ile	Val	394
GCT	GCA	GCC	GTA	CTG	TCT	GGA	AAT	AGA	AAT	TTT	GAG	GGT	CGT	GTA	CAC	CCT	TTG	1242
Ala	Ala	Ala	Val	Leu	Ser	Gly	Asn	Arg	Asn	Phe	Glu	Gly	Arg	Val	His	Pro	Leu	412
ACG	AGG	GCC	AAC	TAC	CTT	GCT	TCT	CCT	CCT	CTT	GTT	GTT	GCT	TAT	GCT	CTT	GCT	1296
Thr	Arg	Ala	Asn	Tyr	Leu	Ala	Ser	Pro	Pro	Leu	Val	Val	Ala	Tyr	Ala	Leu	Ala	430
GGC	ACA	GTG	GAT	ATC	GAT	TTC	GAA	AGT	GAG	CCT	ATT	GGG	GTG	GGA	AAG	GAT	GGA	1350
Gly	Thr	Val	Asp	Ile	Asp	Phe	Glu	Ser	Glu	Pro	Ile	Gly	Val	Gly	Lys	Asp	Gly	448
AAG	AAA	GTA	TTC	TTT	AGG	GAC	ATT	TGG	CCA	ACT	AGC	GAA	GAA	GTT	GCT	GTT	GTT	1404
Lys	Lys	Val	Phe	Phe	Arg	Asp	Ile	Trp	Pro	Thr	Ser	Glu	Glu	Val	Ala	Val	Val	466
GTA	AAC	TCA	AAT	GTG	TTG	CCT	GAC	ATG	TTT	CGT	GCC	ACA	TAC	CAA	GCG	ATC	ACG	1458
Val	Asn	Ser	Asn	Val	Leu	Pro	Asp	Met	Phe	Arg	Ala	Thr	Tyr	Gln	Ala	Ile	Thr	484
GAA	GGA	AAT	GCA	ACT	TGG	AAT	CTT	TTA	TCG	GTT	CCT	GAA	GGA	ACA	CTT	TAC	TCC	1512
Glu	Gly	Asn	Ala	Thr	Trp	Asn	Leu	Leu	Ser	Val	Pro	Glu	Gly	Thr	Leu	Tyr	Ser	502
TGG	GAC	CCA	ACA	TCG	ACG	TAC	ATT	CAC	GAG	CCT	CCG	TAT	TTC	AAA	GAT	ATG	AGC	1566
Trp	Asp	Pro	Thr	Ser	Thr	Tyr	Ile	His	Glu	Pro	Pro	Tyr	Phe	Lys	Asp	Met	Ser	520
ATG	TCT	CCT	CCA	GGG	CCT	CAT	GGG	GTC	AAA	AAC	GCA	TAC	TGC	TTG	CTC	AAT	TTT	1620
Met	Ser	Pro	Pro	Gly	Pro	His	Gly	Val	Lys	Asn	Ala	Tyr	Cys	Leu	Leu	Asn	Phe	538
GGT	GAT	AGT	ATT	ACA	ACT	GAT	CAC	ATC	TCA	CCT	GCT	GGT	AGC	ATC	CAT	AAA	GAT	1674
Gly	Asp	Ser	Ile	Thr	Thr	Asp	His	Ile	Ser	Pro	Ala	Gly	Ser	Ile	His	Lys	Asp	556
AGT	CCT	GCG	GCC	AAG	TAC	CTT	CTC	GAA	CGT	GGG	GTT	GAT	AGA	AGA	GAT	TTC	AAC	1728
Ser	Pro	Ala	Ala	Lys	Tyr	Leu	Leu	Glu	Arg	Gly	Val	Asp	Arg	Arg	Asp	Phe	Asn	574
TCT	TAC	GGA	GTC	GCC	GTG	GTA	ATG	ATG	AGA	TTA	TGG	CAC	GTG	CAC	TTT	GCC	AAC	1782
Ser	Tyr	Gly	Val	Ala	Val	Val	Met	Met	Arg	Leu	Trp	His	Val	His	Phe	Ala	Asn	592
ATT	CGC	ATA	GTC	AAT	AAA	CTA	TTG	AAG	GGG	GAA	GTT	GGA	CCC	AAG	ACG	ATT	CAC	1836
Ile	Arg	Ile	Val	Asn	Lys	Leu	Leu	Lys	Gly	Glu	Val	Gly	Pro	Lys	Thr	Ile	His	610
ATC	CCC	AGT	CGG	GAG	AAA	CTT	TCA	GTA	TTT	GAT	GCT	GCA	ATG	AGA	TAC	AAG	AGC	1890
Ile	Pro	Ser	Arg	Glu	Lys	Leu	Ser	Val	Phe	Asp	Ala	Ala	Met	Arg	Tyr	Lys	Ser	628
GAG	GGG	CAA	GAT	ACA	ATC	ATT	CTC	GCT	GGT	GCC	GAG	TAT	GGA	ATT	GGA	AGT	TCT	1944
Glu	Gly	Gln	Asp	Thr	Ile	Ile	Leu	Ala	Gly	Ala	Glu	Tyr	Gly	Ile	Gly	Ser	Ser	646
CGT	GAT	TGG	GCT	GCC	AAG	GGT	CCA	ATG	CTT	CTG	GGT	GTT	AAA	GCG	GTT	ATA	GCA	1998
Arg	Asp	Trp	Ala	Ala	Lys	Gly	Pro	Met	Leu	Leu	Gly	Val	Lys	Ala	Val	Ile	Ala	664
AAG	AGC	TTT	GAA	CGT	ATA	CAC	CGA	AGC	AAC	TTG	GTC	GGT	ATG	GGT	ATC	ATT	CCT	2052
Lys	Ser	Phe	Glu	Arg	Ile	His	Arg	Ser	Asn	Leu	Val	Gly	Met	Gly	Ile	Ile	Pro	682
CTA	TGT	TTC	AAG	GCT	GGG	GAG	GAT	GCC	GAT	TCT	CTG	GGA	TTG	ACT	GGG	CAC	GAA	2106
Leu	Cys	Phe	Lys	Ala	Gly	Glu	Asp	Ala	Asp	Ser	Leu	Gly	Leu	Thr	Gly	His	Glu	700
CGA	TTC	ACC	ATC	GAT	CTT	CCA	AGC	AAT	GTG	GGT	GAA	ATC	AGA	CCT	GGT	CAG	GAT	2160
Arg	Phe	Thr	Ile	Asp	Leu	Pro	Ser	Asn	Val	Gly	Glu	Ile	Arg	Pro	Gly	Gln	Asp	718
GTG	GCT	GTG	GTG	ACA	GAC	ACC	GGA	AAG	TCT	TTC	AGT	TGC	ATT	CTA	AGA	TTT	GAT	2214
Val	Ala	Val	Val	Thr	Asp	Thr	Gly	Lys	Ser	Phe	Ser	Cys	Ile	Leu	Arg	Phe	Asp	736
ACA	GAG	GTG	GAA	CTG	GCA	TAC	TTC	GAT	CAT	GGT	GGA	ATT	CTG	CAG	TAT	GTC	ATC	2268
Thr	Glu	Val	Glu	Leu	Ala	Tyr	Phe	Asp	His	Gly	Gly	Ile	Leu	Gln	Tyr	Val	Ile	754

FIGURE 4 (iii)

[illegible]

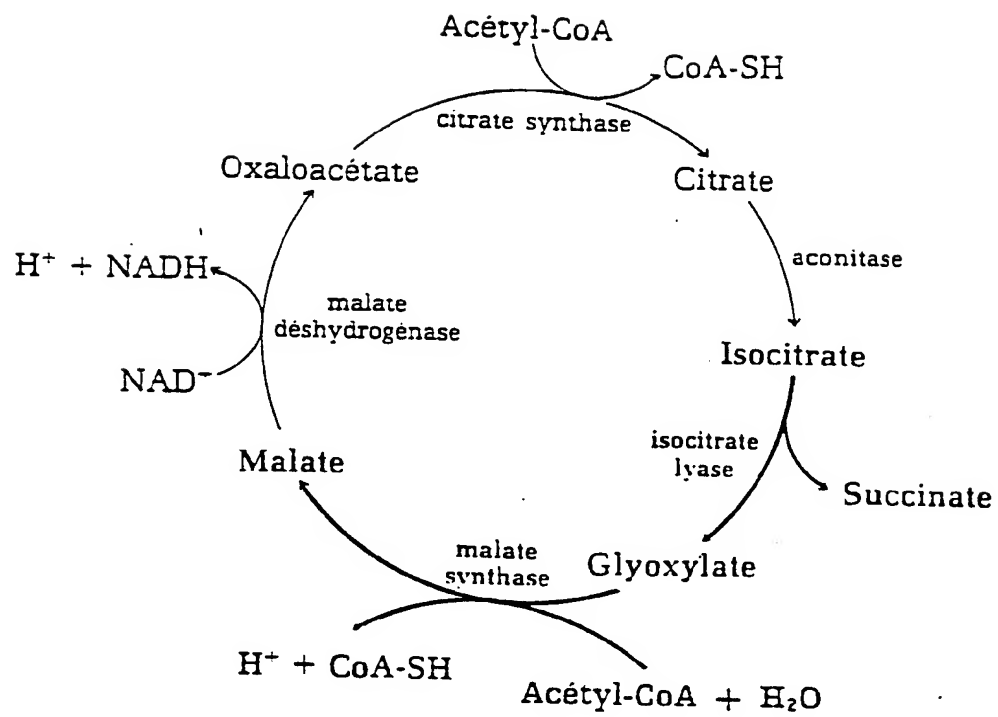
15/36

FIGURE 5



16/36

FIGURE 6



17/36

FIGURE 7

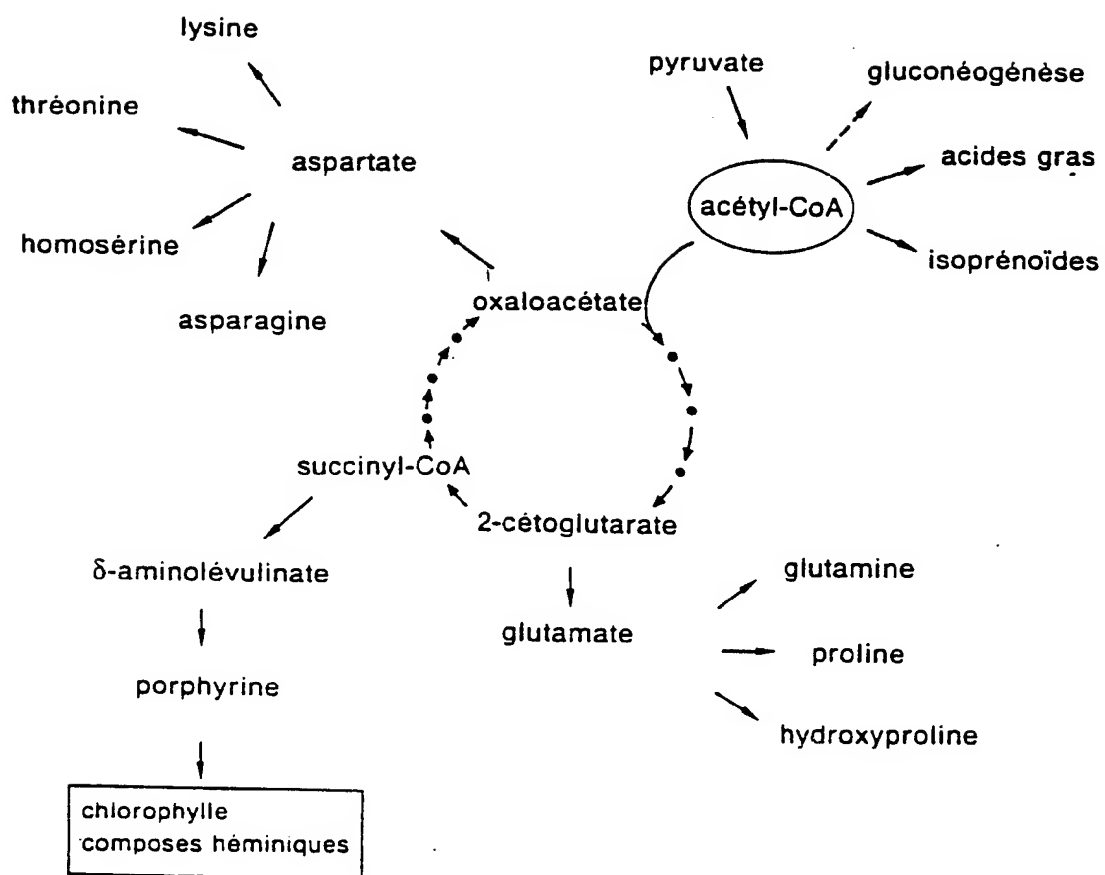


FIGURE 8

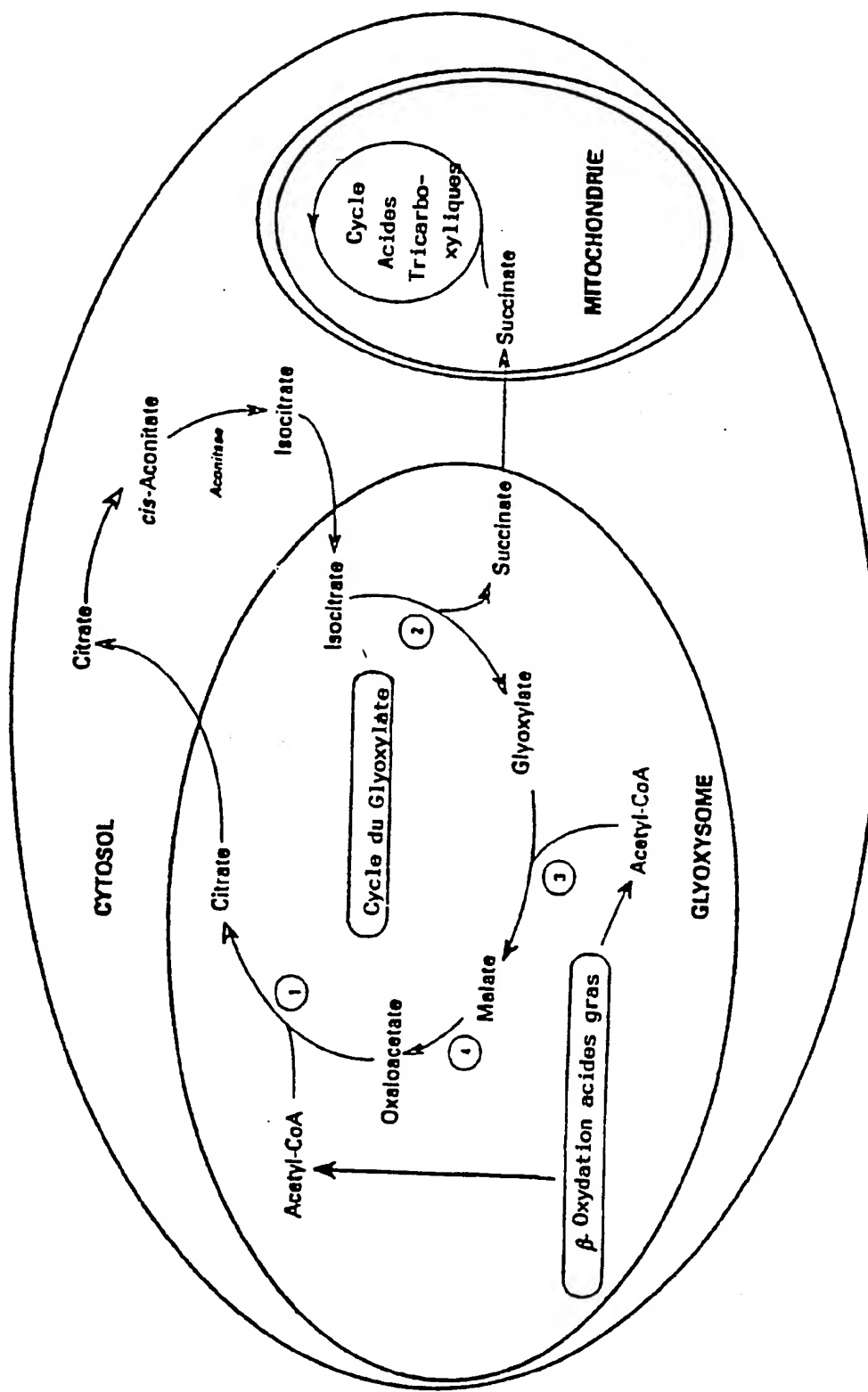


FIGURE 9

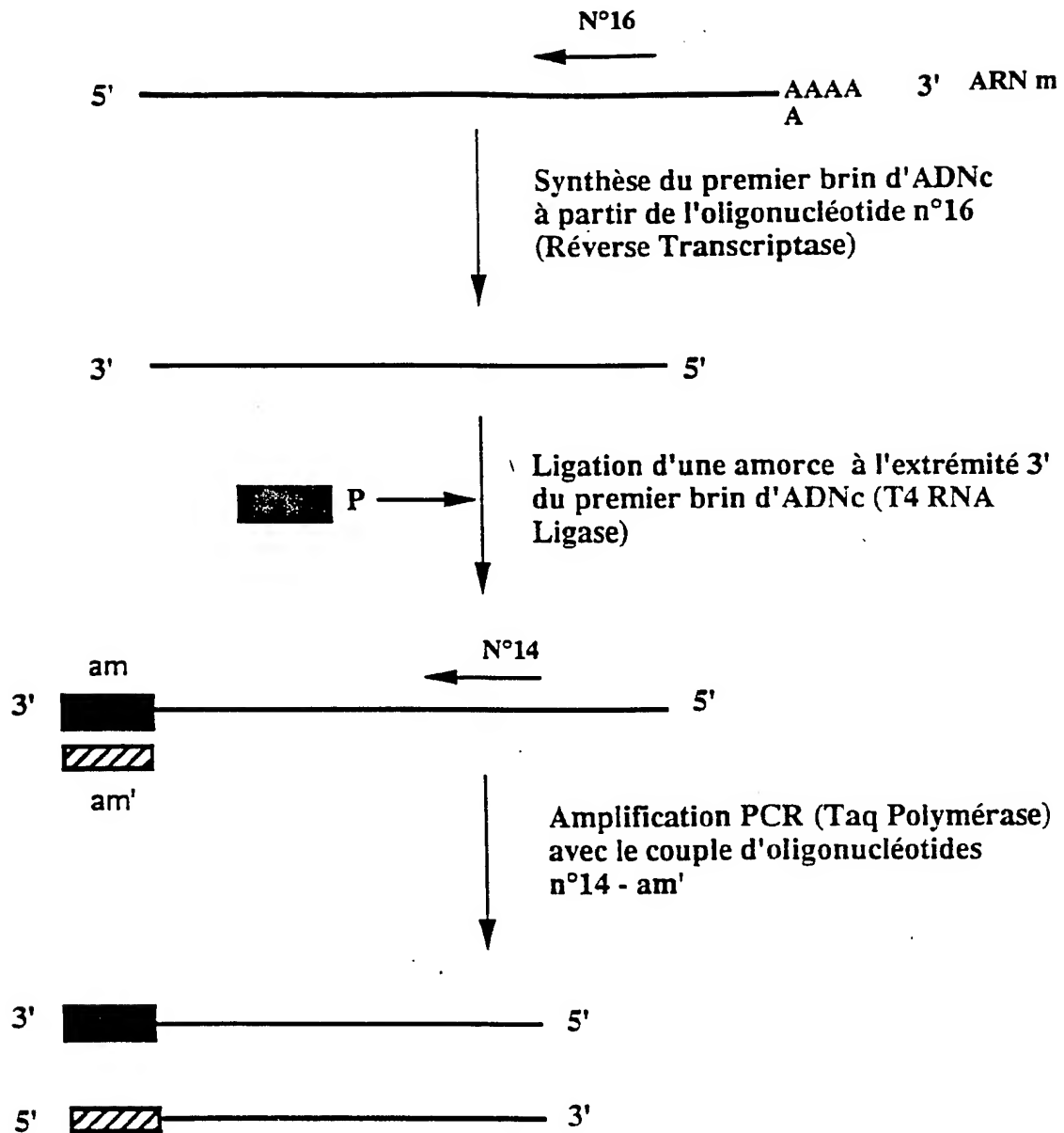
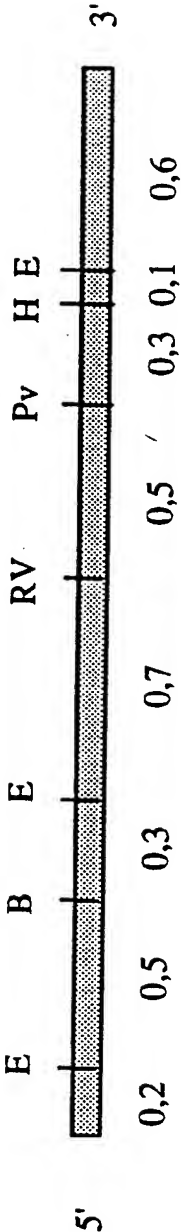


FIGURE 10



21/36

FIGURE 11(a)

		a		-1
A	AAC TTT CTC TCT TCT GAA AAC AGA TCT CTC TAC TTT GCT TCT TCA CTC GAC TTG			54
	Asn Phe Leu Ser Ser Glu Asn Arg Ser Leu Tyr Phe Ala Ser Ser Leu Asp Leu			17
	TAT CTA TCA TCC ATG GCT TCC GAG AAT CCT TTC CGA AGC ATA TTG AAG GCG TTA			108
	Tyr Leu Ser Ser Met Ala Ser Glu Asn Pro Phe Arg Ser Ile Leu Lys Ala Leu			35
	GAG AAG CCT GAT GGT GGT GAA TTC GGT AAC TAC TAC AGC TTA CCT GCT TTG AAC			162
	Glu Lys Pro Asp Gly Gly Glu Phe Gly Asn Tyr Tyr Ser Leu Pro Ala Leu Asn			53
	GAT CCC AGG ATC GAT AAA CTA CCT TAT TCC ATT AGG ATA CTT CTT GAA TCG GCC			216
	Asp Pro Arg Ile Asp Lys Leu Pro Tyr Ser Ile Arg Ile Leu Glu Ser Ala			71
	ATA CGT AAC TGT GAT GAG TTC CAA GTT AAG AGC AAA GAT GTT GAG AAG ATT CTT			270
	Ile Arg Asn Cys Asp Glu Phe Gln Val Lys Ser Lys Asp Val Glu Lys Ile Leu			89
	GAT TGG GAG AAT ACT TCT CCC AAG CAG GTT GAG ATT CCG TTC AAG CCT GCT CGG			324
	Asp Trp Glu Asn Thr Ser Pro Lys Gln Val Glu Ile Pro Phe Lys Pro Ala Arg			107
	GTT CTT CTT CAG GAC TTT ACT GGT GTT CCT GCT GGT GTT GAT CTT GCT TGC ATG			378
	Val Leu Leu Gln Asp Phe Thr Gly Val Pro Ala Val Val Asp Leu Ala Cys Met			125
	AGA GAT GCC ATG AAT AAT CTC GGT GGT GAT TCT AAT AAA ATT AAT CCG CTG GTC			432
	Arg Asp Ala Met Asn Asn Leu Gly Gly Asp Ser Asn Lys Ile Asn Pro Leu Val			143

22/36

FIGURE 11(b)

A. thaliana : Ser-Ser-Met-Ala-Ser-Glu-Asn-Pro-Phe-Arg-Ser-Ile-Leu-Lys

B C. melo : Ser-Ser-Met-Ala-Ala-Glu-Asn-Pro-Phe-Lys-Glu-Asn-Leu-Thr

S. tuberosum : ? - Thr-Met-Ala-Ala-Glu-Asn-Pro-Phe-Lys-Gly-Ile-Leu-Thr

FIGURE 12

séquences	arabaco	melonaco	acon_ecoll	acon_plg	acon_yeast	lreb_human	lreb_mouse	lreb_rabbit
arabaco	100 (0)							
melonaco	86.8 (5.0)	100 (0)						
acon_ecoll	53.6 (12.8)	53.7 (13.1)	100 (0)					
acon_plg	30.2 (13.2)	14.4 (13.1)	27.5 (15.4)	100 (0)				
acon_yeast	30.2 (14.5)	15.6 (13.9)	27.1 (14.9)	65.1 (9.1)	100 (0)			
lreb_human	58.2 (11.9)	57.8 (12.3)	50.5 (13.6)	26.6 (14.7)	27.5 (13.7)	100 (0)		
lreb_mouse	59.5 (11.5)	59.1 (12.2)	51.5 (12.8)	29.1 (14.7)	29.0 (14.8)	88.9 (3.5)	100 (0)	
lreb_rabbit	59.5 (11.8)	59.9 (11.1)	52.3 (12.4)	28.7 (15.2)	29.4 (14.4)	89.5 (2.9)	92.6 (2.7)	100 (0)

FIGURE 13

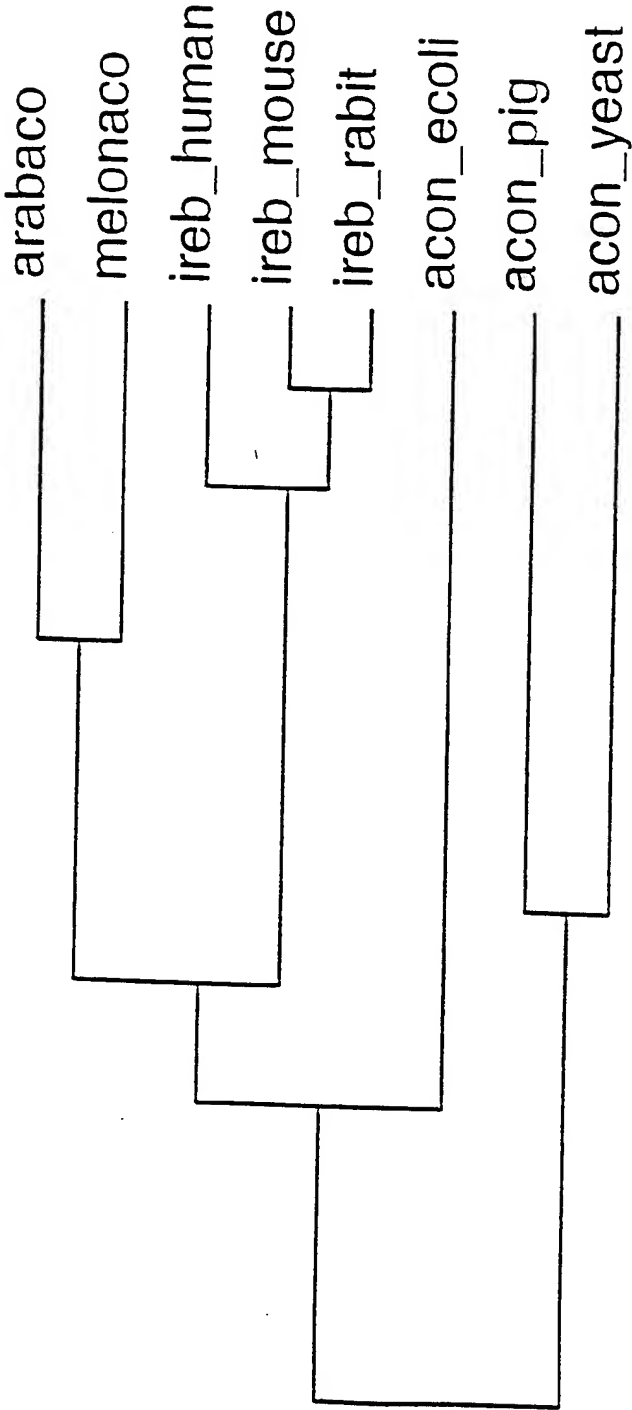


FIGURE 14

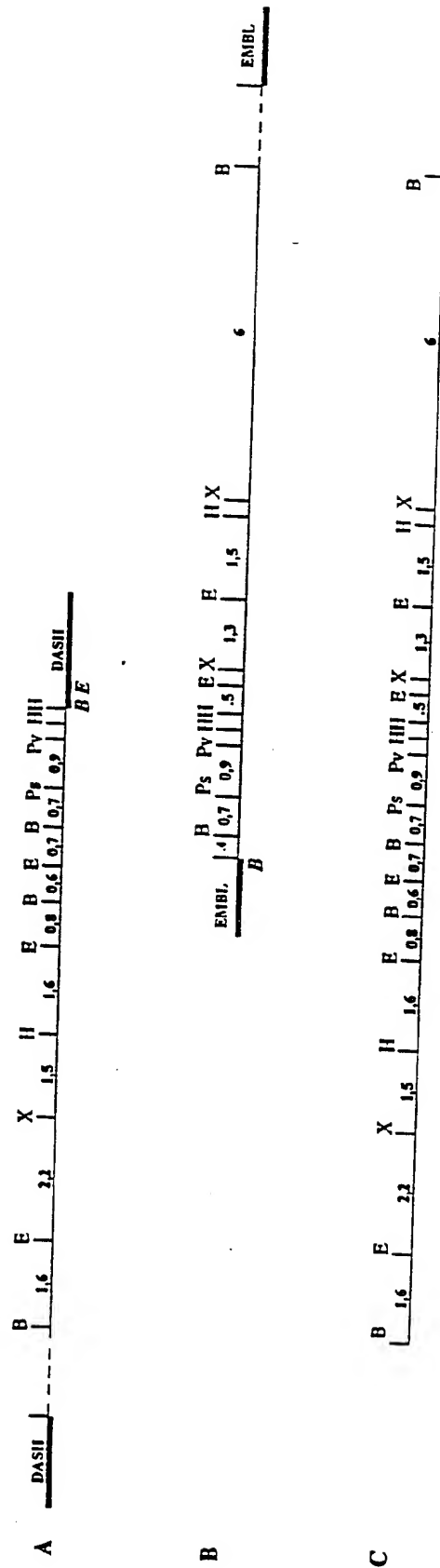


FIGURE 15

	C	
A		C
C		A
C		C
	CG	
	CG	
	AU	A
	GC	
A		U
A		
A		
A		
G		
	CG	
	AU	
	GC	
	UA	
	CG	
	UA	

FIGURE 16

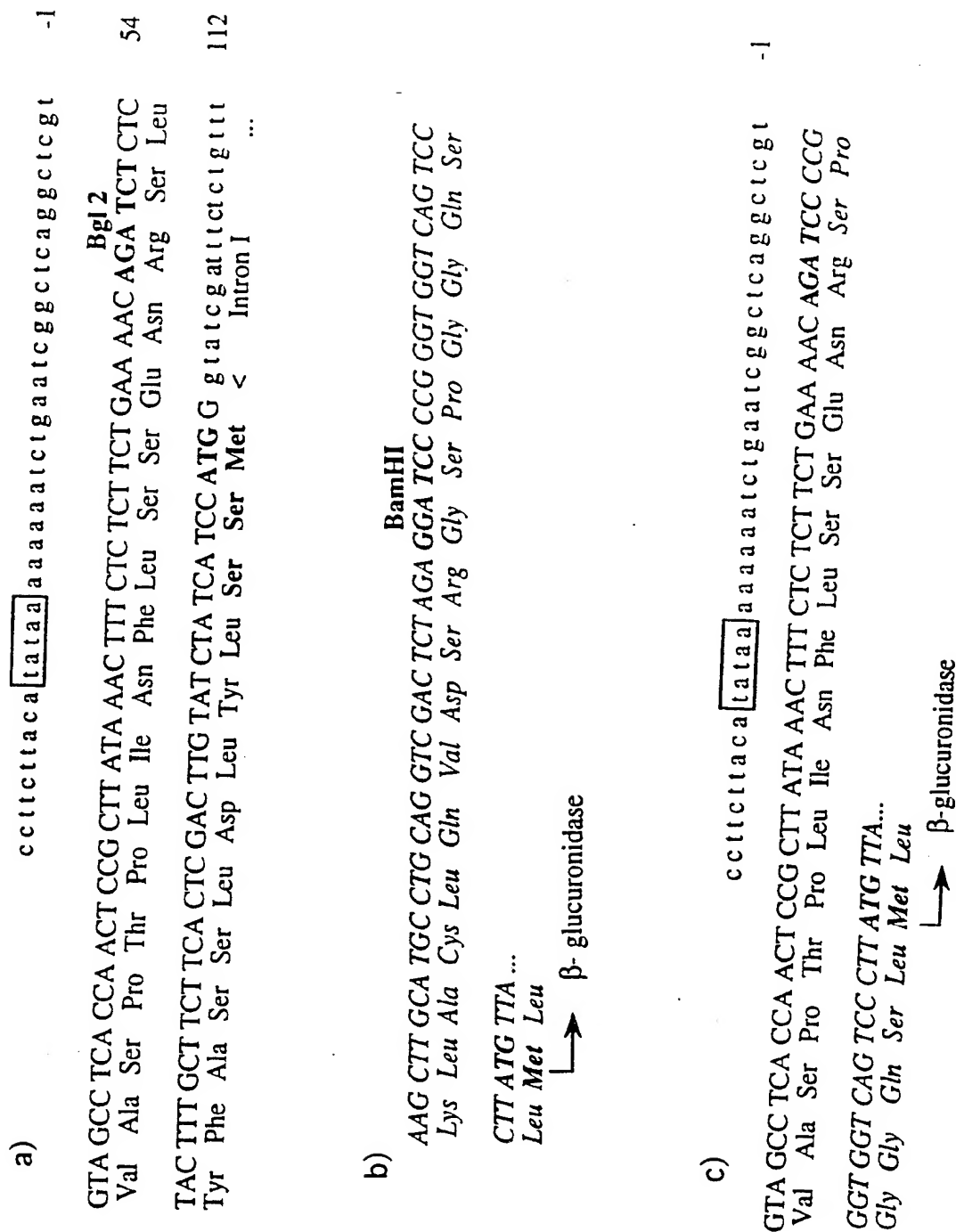


FIGURE 18(a)

29/36

10 20 30 40 50 60 70 80 90
 CAGTGATATC TGCTGGTAGT GTTTACGTAC TATAAGGATT GCGGATTATG TTCATGTGAT TAGCAGTTAG TGGAAATTTT TTGAGATGGA
 100 110 120 130 140 150 160 170 180
 GGCTCCGACC CCACCTCCGT TAACAATATA ACGGAGATCA GCTAATTTGG CTTTGATACA ACATTATGGC TATTTCTGAA CCAACCAAGC
 190 200 210 220 230 240 250 260 270
 CATTACTATG GCTACTAAGCT GTTGCCGGGA CCTTAGTGCT AAGTAGGCTA TAAGGTTGTA GTGCATGATC CTATTGACCT ATATTTTCTAGA
 280 290 300 310 320 330 340 350 360
 CTGTGTTGTA CGGGGTGATT TTGCGCTGTT GTTAAATTTAT CATCTTGCTT TCCCTATTCC CATATGGATA GGTTAAGGAA TCCTATAATA
 370 380 390 400 410 420 430 440 450
 TAATTGACTA ATTTTCATGT GCGGACCAAC TTGATGACTC GTTTAGTTGA CTTGAAATGT TCTGCACCTC AGGATAAATT TAACTTCGGT
 460 470 480 490 500 510 520 530 540
 GAGATCTGTT AATGCAGAGG CATTTCCTTG TTCCAGTAGC TAATCATGTT GCCTTACTTT GTGTGCAGCT ACTGTGCATG CCTTCAAGGA
 550 560 570 580 590 600 610
 T ATT CTG ACC AGC CTC CCT AAG CCT GGA GGT GGT GAA TAT GGA AAG TTC TAC AGC CTT CCT GCA CTA AAT GAT
 I L T S L P K P G G G E Y G K F Y S L P A L N D>
 620 630 640 650 660 670 680 690 700
 CCA AGG ATT G GTAATTG TGCTGCTTCC TTTTGAAATT GTTTTCTCTT TACTTCAGCA GTTGCTAGGT TACATTGCCA ACGGGTGTGT
 P R I D>
 710 720 730 740 750 760 770 780 790
 ATGTTTCACTG TGATTTTAGT ATCCTAAACT GAAGAAAAAT TATAATACAT GCTTGTAATA TTACTTCACCT TAAGTTCCCC ACTTCCAGTC
 800 810 820 830 840 850 860 870 880
 AAAACGAGGT TTTATTAGTT GTTTTCTATT CTCAATAGTC AATAATCCAT GTTAAGTGAA AATGAAATAT TGACACAATT CTATGGCATT
 890 900 910 920 930 940 950 960 970
 CCTTCCTTAA GAAAGCTATA TATTCACACA TCAAAATTGG ACTGGTTAAT TAACTTATA CAGAGAACAT ATCACACAAT TCTATGGCAT
 980 990 1000 1010 1020 1030 1040 1050 1060
 ACCTTCCTTA AGAAAGCTAT ATATAATATA TTCTCTAGTC ATACCTTTTG TGTTCCTATC AATTCTTTAG TATGTGACAC AGCAAAATGT
 1070 1080 1090 1100 1110 1120 1130 1140
 ACATAGTTAT ATCTTCTATT GTTGTAAGTC AATGTATATT ATATCTTTTG TTATGTATCA G AT AAG CTC CCG TAC TCC CTC CGT
 K L P Y S V R>
 1150 1160 1170 1180 1190 1200 1210
 ATT CTT CTT GAG TCA GCT ATC CGT AAC TGT GAT AAC TTC CAA GTT ACC AAG AAT GAT GTT GAG AAA ATA ATT GAC
 I L L E S A I R N C D N F Q V T K N D V E K I I D>
 1220 1230 1240 1250 1260 1270 1280 1290
 TGG GAA AAC ACA TCT CCA AAG CTG GCC GAG ATA ECA TTC AAG CCA GCT CGA GTT CTT CAG GTTTGAAC
 W E N T S P K L A E I P F K P A R V L L Q>
 1300 1310 1320 1330 1340 1350 1360 1370
 TGATAACATC ACACATGATC TCCTTTGTTT CTAAATGCCC TTGCTCATC CATTGACATG TACTTTTAG GAC TTC ACT GGA GTT CCA
 D F T G V P>
 1380 1390 1400 1410 1420 1430 1440 1450
 GCA GTT GTT GAT CTT GCA GCT ATG CGT GAT CCG ATG GCC AAG TTG GGC AGT GAT GCG AAC AAG ATT AAC CCA TTG
 A V V D L A A M R D A M A K L G S D A N K I N P L>
 1460 1470 1480 1490 1500 1510 1520 1530 1540
 GTACATTC ATCTTTTCTC GTTGTATGT GAGATCTGTA AATTTTCATCA ATTTGTGCTC AACCACGTGT TATTTCCGGA GAAAAGATGA
 1550 1560 1570 1580 1590 1600 1610 1620 1630
 GTTGACATAG GGCAGAATAC CAGAAGTTTA TATGCTCTAT ATTTCTGGTG GACACCTATC CATGTCTTGG GCAATCTTTT TTCTCTTGT
 1640 1650 1660 1670 1680 1690 1700
 GCACATCCTT GTTTGTACTG CAG GTC CCG GTG GAT CTT GTT ATT GAT CAC TCA GTG CAG GTG GAT GTC GCA AGG TCG
 V P V D L V I D H S V Q V D V A R S>

FIGURE 18(b)

30/36

1710 1720 1730 1740 1750 1760 1770 1780
 CAG AAT GCC GTA CAA GCA AAC ATG CAA CTG CAA TTT ACT CGC AAC AAG GAA AGA TTT GGT TTC CTT AAA TGG GGC
 Q N A V Q A N M E L E F S R N K E R F G F L K W G>

1790 1800 1810 1820 1830 1840 1850 1860
 TCT AGT GCT TTC CAA AAT ATG CTA GTT GTT CCT CCT GGT TCT GGT ATT GTG CAC CAG G TAAATGCTCT TGACTTGCTG
 S S A F Q N M L V V P P G S G I V H Q>

1870 1880 1890 1900 1910 1920 1930 1940
 AAAATTCCTG TAGTGCTACC ATGTCTTCTA TGTGTGTAAT TCCGCCTTTG TTCTTTGTGT ACAAATATCA G GTC AAT CTT GAG TAT
 V N L E Y>

1950 1960 1970 1980 1990 2000 2010 2020
 CTT GGC AGA GTT GTT TTC AAC ACA GAT GGC ATT CTC TAT CCT GAC AGC GTA GTT GGT ACT GAT TCA CAC ACC ACT
 L G R V V F N T D G I L Y P D S V V G T D S H T T>

2030 2040 2050 2060 2070 2080 2090
 ATG ATT GAT GGT TTG GGT GTT GCT GGC TGG GGG GTT GGT GGT ATA GAA GCG GAG GCT ACA ATG CTT GGT CAG
 M I D G L G V A G W G V G G I E A T M L G Q>

2100 2110 2120 2130 2140 2150 2160 2170 2180
 GTAAATC TCATGTGGG CCACTCATCC TTAATGACA GATTCTAAT CATTACCTTT TTTTGTGTTT TAAGGAGAAA ATTGCTCTTC

2190 2200 2210 2220 2230 2240 2250 2260 2270
 TGACATTTTT CCCACACCCC AAATTTTGA ATGGCTTATT TACTTTGTGC CCTTTTATTT GTTTGATTAT ACATGGGATT TAGCTAAACA

2280 2290 2300 2310 2320 2330 2340 2350
 TCAAACATCAT TATTAATTGG TTTTAGGCTG GTTCACCTGA CCTGTTTGTGTT ATTGCCTTAT GCAG CCA ATG AGC ATG GTT TTG CCT
 P M S M V L P>

2360 2370 2380 2390 2400 2410 2420 2430
 GGT GTT GTT GGC TTC AAG TTA ACT GGG AAG TTA CGG AGT GGT GTC ACG GCC ACT GAT CTT GTT CTT ACA GTA ACA
 G V V G F K L T G K L R S G V T A T D L V L T V T>

2440 2450 2460 2470 2480 2490 2500
 CAA ATG CTA AGG AAA CAT GGT GTT GTT GGC AAA TTT GTT GAA TTC TAT G G TAAGTCCTTG ATGCTAGTAG
 Q M L R K H G V V G K F V E F Y G>

2510 2520 2530 2540 2550 2560 2570 2580
 ATGGAATGTT CTGTCTGAT GCAATTAAAC TATATGTTAT TGACATATTT CCCTTTGTGT TCTCTATCCA G GT GAA GGT ATG GGT
 E G M G>

2590 2600 2610 2620 2630 2640 2650 2660
 AAG CTG TCT TTA GCT GAC AGG GCC ACA ATT GCG AAC ATG TCA CCA GAA TAT GGA GCT ACC ATG GGC TTC AAC CCT
 K L S L A D R A T I A N M S P E Y G A T M G F N P>

2670 2680 2690 2700 2710 2720 2730
 GTG GAC CAT GTG ACA TTG GAT TAC CTT AAA CTG ACA GGC CGC AGT GAT GAA ACT GTATGT TTTAGTACCT
 V D H V T L D Y L K L T G R S D E T>

2740 2750 2760 2770 2780 2790 2800 2810 2820
 GCAAATGCAA ATTATTTTTC GTTCTATTG CGATGGTAGT TTGTGGCAAA CATACTCATA AAATATTTGT CAATGCATAC ATATCTCCCC

2830 2840 2850 2860 2870 2880 2890 2900 2910
 CTGTGGGGAT GTGTGAATTC AAGCTTCTCT ACTGCGTTTC TTCAGCTAGG ATCTGTTTCCA TATTGATTC CGAAAAGTGA CTCACTAAAC

2920 2930 2940 2950 2960 2970 2980 2990 3000
 TTCAAGGACA CAAAGTCGTC TCTCCCAAAA TTCCATGTCC TTATTGCCAT CTATACTACC ATTTCCATTC AGCGTCTTAA TTTCGTATAT

3010 3020 3030 3040 3050 3060 3070 3080
 AGAACATCAT TTTGATTTTT GTGAGAAGAA TTCTGCATGC ACTCCATTTT ATAAGTTAAA CCTGTTTATT TAG GTG TCA ATG ATT
 V S M I>

3090 3100 3110 3120 3130 3140 3150 3160
 GAA GCA TAC CTG CGA GCC AAT AAA ATG TTT GTG GGC TAC AAT GAG GTAGATTAAA TGAGGTAGAT TTCTTTACTC
 E A Y L R A N K M F V G Y N E>

3170 3180 3190 3200 3210 3220 3230 3240
 ATTGCAGCAC TTGAGTCCCA GATATTATTT TGTGTGCAAA ATATAAGCAT ATTTGCATAT GTTGGCAG CCT CCA ACA GAG AGA ATA
 P P T E R I>

FIGURE 18(c)

31/36

3250 3260 3270 3280 3290 3300 3310 3320
TAC TCA TCT TAT CTT GAG CTG AAC CTT GAT GAG GTG GAG CCT AGC ATG TCT GGC CCA AAG AG GT ACGATCCTGA
Y S S Y L E L N L D E V E P S M S G P K R>

3330 3340 3350 3360 3370 3380 3390 3400
GACTCTTCTG ATTGGAAGTT ATCGCACTTC ATATTTTATG GGGTGGGGTG GGCTGTGATA TTGGGCTTAT GOTTCTGAAC TTGTAG G

3410 3420 3430 3440 3450 3460 3470 3480
CCT CAT GAT CGT GTC CCT TTG AAG GAA ATG AAA TCA GAT TGG CAT GCT TGC CTG GAC AAC AAA GTT GGG TTC AAG
P H D R V P L K E M K S D W H A C L D N K V G F K>

3490 3500 3510 3520 3530 3540 3550 3560 3570
GTAAGTGC TTAGTGTAT ATCATATGTT TTTGCATAAC ATTGCTTCA TATGATGTTA CTTACTCGAT ACATGTATGA GTTTCAG GGT
G>

3580 3590 3600 3610 3620 3630 3640
TTT GCA GTG CCA AAG GAG CAG CAG GAT AAG GTT GTC AAA TTT GAC TTT CAT GGG CAG CCA GCA GAA ATG AAG CAT
F A V P K E Q Q D K V K F D F H G Q P A E M K H>

3650 3660 3670 3680 3690 3700 3710 3720
GGT AGT GTT GTT ATA GCA GCA ATA ACT AGC TGC ACA AAC ACC TCA AAT CCT AGT GTT ATG CTT GGT GCT GGT CTT
G S V V I A A I T S C T N T S N P S V M L G A G L>

3730 3740 3750 3760 3770 3780 3790 3800
GTA GCT AAG AAA GCT TGC GAG TTG GGT CTT GAC GTTAGT CCTCTGTAA TGCAATCATG GGCTGTTGTG CTGAGTGCCT
V A K K A C E L G L E>

3810 3820 3830 3840 3850 3860 3870 3880
TTCTGTGCTT TCTGTTGATG TGAGATAGCA CTTCTTTTCA G GTT AAA CCA TGG GTG AAG ACA AGC CTT GCC CCT GGA TCA
V K P W V K T S L A P G S>

3890 3900 3910 3920 3930 3940 3950 3960
GGA GTT GTC ACG AAG TAC TTG CTT CAG AG G TATTGGCAT TGTTTAAATG CGTATCAGAT CAGGAGTGT CAGGAGTTAT
G V V T K Y L L Q S>

3970 3980 3990 4000 4010 4020 4030 4040
CATTGAAACG TAGAATTCA ATGTGTATCA GATGTCATTT AATTGAATCT GGTGATTCTT AATTACAG T GGT CTT CAA GAA TAC
G L Q E Y>

4050 4060 4070 4080 4090 4100 4110 4120
CTC AAC CAG CAA GGA TTC CAT ATC GTT GGC TAT GGC TGT ACC ACT TGC ATT GGA AAC TCT GGT GAT TTC GAT GAA
L N Q Q G F H I V G Y G C T T C I G N S G D F D E>

4130 4140 4150 4160 4170 4180 4190 4200
TCT GTA TCA ACT GCC ATT ACA GAA AAT G GT AAGTTAATTA TACGATGATT GACTATTAG CTGTAACAAA TTTCAGTTT
S V S T A I T E N D>

4210 4220 4230 4240 4250 4260 4270 4280
TTTCTTTATG ATTGGCTTAT TGCTCAGCTC TTTTGCACAG AT GTT GTT GCT GCT GCT GTT TTG TCG GGT AAC CGT AAC TTC
V V A A A V L S G N R N F>

4290 4300 4310 4320 4330 4340 4350
GAG GGT CGT GTC CAC CCT TTA ACT CGG GCT AAC TAT CTT GCT TCG CCA CCT CTT GTT ATT CGT TAT CGA CTT GCC
E G R V H P L T R A N Y L A S P P L V I R Y R L A>

4360 4370 4380 4390 4400 4410 4420 4430 4440
GGA ACT GTAAATAC TTTTCTCTCA CAACTATTG TAGGGCTGCA TGCTCAATAA ACTTCAGAAC ATCTGTGTTT CTGGTTTCAT
G T>

4450 4460 4470 4480 4490 4500 4510
GGAAGTAGTT AATATTGAA GTTGTACTTT CAG GTT GAT ATT GAT TTT GAG AAA GAG CCC ATT GGA TTT GGA AAG GAT
V D I D F E K E P I G F G K D>

4520 4530 4540 4550 4560 4570 4580 4590
GGC AAG GAA GTC TAC TTC AGG GAT ATA TCG CCC TCG ACA GAA GAA ATT GCT GAG GTATATTC TGTAATCAAA
G K E V Y F R D I W P S T E I A E>

4600 4610 4620 4630 4640 4650 4660 4670 4680
CTTGTGTGTA TTTCAGTTAC TATAGTACAT TCATGTATAC CAGAT---AA TTATCTTGTA GCAGTAATAA TTGTATTCA AATGCTAGCT

4690 4700 4710 4720 4730 4740 4750

FIGURE 18(d)

32/36

TACAG GTT GTC CAA TCG AGT GTG CTG CCT GAC ATG TTC AAG GGC ACC TAT GAG GCT ATC ACA AAA GGC AAC CCA
 V V Q S S V L P D M F K G T Y E A I T K G N P>

4760 4770 4780 4790 4800 4810 4820
 ATG TGG AAC CAG CTA ACT GTC CGA GAA GGA TCA CTC TAC TCA TGG GAT TCC AAA TCC ACT TAC TCC ATG AGG AAG
 M W N Q L T V R E G S L Y S W D S K S T Y S M R K>

4830 4840 4850 4860 4870 4880 4890 4900
 CTT ACT TTA AGG AAC ATG ACC ATG TCC CCA CCT GCC CCG TCT ACA GTG AAA GAT GCC TAC TGC TTG CTG AAC TTC
 L T L R N M T M S P P A P S T V K D A Y C L L N F>

4910 4920 4930 4940 4950 4960 4970
 GGG GAC AGC ATT ACT ACA GAT CAC ATT TCA CCT GCA GGA AGC ATA CAC AAA GAC AGT CCT GCT GCC AAG TAT TTG
 G D S I T T D H I S P A G S I H K D S P A A K Y L>

4980 4990 5000 5010 5020 5030 5040 5050
 ATG GAG CGT GGT GTG GAC CGG AAG GAC TTC AAC TCA TAT GGT AGC CGT CGT GGT AAT GAT GAA GTA ATG GCA AGG
 M E R G V D R K D F N S Y G S R R G N D E V M A R>

5060 5070 5080 5090 5100 5110 5120
 GGA ACG TTT GCA AAC ATT AGG ATC GTG AAC AAG TTT TTG AAC GGA GAA GTT GGA CCC AAG ACC ATT CAT GTT CCT
 G T F A N I R I V N K F L N G E V G P K T I H V P>

5130 5140 5150 5160 5170 5180 5190 5200 5210
 ACT GGG GAG AAG CTT TCT GTT TTT GAT GCG GCC ATG GTATG TATTGTGAGG TTTTATTCCT TTGCTTCTA TCAACGTGAC
 T G E K L S V F D A A M>

5220 5230 5240 5250 5260 5270 5280
 CTCCTAATAC CCATGCTCCG TACTGTCAG AGA TAC AAA TCT CAG GGC CAT GCT ACT ATA ATC CTC GCT GGC GCT GAG TAT
 R Y K S Q G H A T I I L A G A E Y>

5290 5300 5310 5320 5330 5340 5350 5360
 GGA AGC GGC AGT TCT CGT GAT TGG GCT GCT AAG GGT CCG ATG CTC CTG GCA AGTTCTATTC AGAGTTTATG
 G S G S S R D W A A K G P M L L>

5370 5380 5390 5400 5410 5420 5430 5440
 AGAATTTTAC CTTGTATCGG AAAACATATA GTATTGATTG ATGAGGTATG CATTGTTTTC AG GGA GTT AAA GCT GTA ATT GCC
 G V K A V I A>

5450 5460 5470 5480 5490 5500 5510
 AAG AGC TTT GAG CGT ATC CAC AGA AGC AAC TTG GTG GGG ATG GGA ATC ATT CCT CTT TGC TTC AAA GCT GGT GAG
 K S F E R I H R S N L V G M G I I P L C F K A G E>

5520 5530 5540 5550 5560 5570 5580 5590
 GAT GCT GAT TCA ATT GGC CTC ACT GGT CAT GAG CGG TAC AGC ATC GAT CTC CCT ACC AAC CTC AGT GAG ATC CGT
 D A D S I G L T G H E R Y S I D L P T N L S E I R>

5600 5610 5620 5630 5640 5650 5660
 CCC GGC CAG GAT GTG ACT GTT ACT ACC GAC AAT GGG AAA TCT TTC ACT TGC ATC GTT CGC TTT GAC ACT GAG
 P G Q D V T V T T D N G K S F T C I V R F D T E>

5670 5680 5690 5700 5710 5720 5730 5740 5750
 GTATA AGCATTTAAG AAGACATCTT TTAACCTTTC TAGCAGGTAT GCCATTTGCA TTGTTATTGA ACTAATCAAT CACCAATTCT

5760 5770 5780 5790 5800 5810 5820
 TGTTTTGTTG TAG GTG GAG CTG GCG TAC TTC AAC CAT GGA GGC ATC CTC CCT TAT GTC ATC CGC AAC TTG GCA GCT
 V E L A Y F N H G G I L P Y V I R N L A A>

5830 5840 5850 5860 5870 5880 5890 5900 5910
 GCG CAT AAC TGAGG CAAGATCACA CACCGTTTGA CCCACAATGG CAGTGGCAGG CTCCATGTTT CCAGGGCTAG CGACCTGGTT
 A H N>

5920 5930 5940 5950 5960 5970 5980 5990 6000
 CTCAATATAG TGACATTTGT GATGGATCTA GGGTCTCTT TTGTCGATTG AGCCCATAAA TAAAGTGTGA AATAAAAGCC GCAGGTGGAG

6010 6020 6030
 GCAGAACCTG CTGATTGAGG GAACAGTTTT GCCTGT

FIGURE 19

33/36

ZMACO	TV-----HAFKDILTSLPKPGGGEYGFYSLPAL
ATACO	ASPTPLINFLSSENRSLYFASSLDLYLSSMASENPFPSILKALEKPDGGEFGNYSLPAL

ZMACO	NDPRIDKLPYSVRILLESAIRNCNDFQVTKNVDEKIIDWENTSPKLAIEIPFKPARVLLQD
ATACO	NDPRIDKLPYSIRILLESAIRNCDEFQVKSVDVEKILDWENTSPKQVEIPFKPARVLLQD

ZMACO	FTGVPAVVDLAAMRDAMAKLGS DANKINPLVPVDLVIDHSVQVDVARSQNAVQANMELEF
ATACO	FTGVPAVVDLACMRDAMNNLGGDSNKINPLVPVDLVIDYSVQVDVARSENAVQANMELEF

ZMACO	SRNKERFGFLKWGSSAFQNLVPPGSGIVHQVNLEYLGRVVFNTDGI LYPDSVVGTD SH
ATACO	QRNKERFAFLKWGSNAFHNMLVPPGSGIVHQVNLEYLARVVFNTNGLLYPDSVVGTD SH

ZMACO	TTMIDGLGVAGWGVGGIEAEATMLGQPM SMVLPGVVGFKL TGKLRSGVTATDLVLTVTQM
ATACO	TTMIDGLGVAGWGVGGIEAERPMLGQPM SMVLPGVVGFKL TGKLRDGMTATDLVLTVTQM

ZMACO	LRKHGVVGKFVEFYGEGMGKLSLADRATIANMSPEYGATMGFNPDHVTLDYLKLTGRSD
ATACO	LRKHGVVGKFVEFHGEGMRELSLADRATIANMSPEYGATMGFFPDHVTLQYLRLTGRSD

ZMACO	ETVSMIEAYLRANKMFVGYNEPPTERIYSSYLELNLDEVEPSMSGPKRPHDRVPLKEMKS
ATACO	DTVSMIEAYLRANKMFVDYSEPESTVYSSCLELNLDEVEPCVSGPKRPHDRVPLKEMKA

ZMACO	DWHACLDNKVGFGKFAVPKEQQDKVVKDFHGO PAEMKHGSVVIAAITSCTNTSNPSVML
ATACO	DWHSCLDNRVGFGKFAVPKEAQSKAVEFNENGTTAQLRHGDVIAAITSCTNTSNPSVML

ZMACO	GAGLVAKKACELGLEVKPWKTS LAPGSGVVT KYLLQSGLOEYLNQQGFHIVGYGCTTCI
ATACO	GAALVAKKACDLGLEVKPWIKTS LAPGSGVVT KYLAKSGLOKYLNLQGF SIVGYGCTTCI

ZMACO	GNSGDFDES VSTA ITENDVVA AVL SGNRNFEGRVHPLTRANYLASPPLVIRYRLAGTVD
ATACO	GNSGDIHEAVASAI VDN DLVASAVLSGNRNFEGRVHPLTRANYLASPPLVVA YALAGTVD

ZMACO	IDFEKEPIGFGKDGKEVYFRDIWPSTEEIAEVVQSSVLPDMFKGTYEAITKGNPMWNQLT
ATACO	IDFETQPIGTGKDGKQIFFRDIWPSNKEVAEVVQSSVLPDMFKATYEAITKGNMWNQLS

ZMACO	VREGSLYSWDSKSTYSMRKLT LRNM TMSPPAPSTVKDAYCLLNFGDSITTDHISPAGSIH
ATACO	VASGTLYEWDPKSTYIHEPPYFKGMTSPPGPHGVKDAYCLLNFGDSITTDHISPAGSIH

ZMACO	KDSPAAYLMERGVDRKDFNSYGSRRGNDEV MAR---GTFANIRIVNKFLNGEVGPKTIH
ATACO	KDSPAAYLMERGVDRRDFNSYGV---AVVMMRLWREHFANIRIVNKLKGEVGPKT VH

ZMACO	VPTGEKLSVFDAAAMYKSQGHATIILAGAEYSGSSRDWAAKGPM LLGVKAVIAKSFERI
ATACO	IPTGEKLSVFDAAMKYRNEGRDTIILAGAEYSGSSRDWAAKGPM LLGVKAVISKSFERI

ZMACO	HRSNLVGMGIIPLCFKAGEDADSIGLTGHERYSIDLPTNLSEIRPGQDVTVT TDNKSFT
ATACO	HRSNLVGMGIIPLCFKAGEDAETLGLTGQELYTIELPNNVSEIKPGQDVTVT TNNKSFT

ZMACO	CIVRFDTEVELAYF NHGGILPYVIRNLAAAHN
ATACO	CTLRFDTEVELAYFDHGGILQYVIRNLI--KQ

36/36

FIGURE 21

DI
(1024)

5' 30 36 3'

M A A E N P F

ccggaattc atg gcg gcg gaa aac ccg tt

t a t a g t a

t t t

c c c

DII
(2048)

5' 194 189 3'

S G W K L F

cccaagctt cct ccc cca ttt caa gaa

tga t c t g a

a a a

g g g

DIII
(256)

5' 268 263 3'

M P Q G L M

cccaagcttc atc ggc tgc ccc agc at

t t t t a

a a a

g g g

DIV
(192)

5' 240 245 3'

H T T M I D

ccggaatt cac acg acg atg ata gac gg

t a a c t

t t t

c c c

DV
(256)

5' 338 344 3'

A T M G F F P

ccggaattc gca acg atg ggg ttc ttc cc

g a a t t

t t t

c c c

DVI
(32)

5' 423 418 3'

H W D A K M

cccaagctt gtg cca gtc cgc ttt cat

a a t c

a

g

INTERNATIONAL SEARCH REPORT

Internal 1 Application No

PCT/EP 95/00263

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/60 C12N15/82 A01H5/00 C12P21/08

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N A01H C12P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	THE BIOCHEMICAL JOURNAL, vol. 294, pages 103-107, COURTOIS-VERNIQUET, F., ET AL. 'Lack of aconitase in glyoxysomes and peroxisomes' see page 104, right column, line 3 - line 15 ---	16
X	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 87, WASHINGTON US, pages 7958-7962, ROUAULT, T.R., ET AL. 'Cloning of the cDNA encoding an RNA regulatory protein - the human iron-responsive element binding-protein' cited in the application see figure 2 --- -/--	4, 5, 10, 11

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

22 June 1995

Date of mailing of the international search report

3.07.95

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+ 31-70) 340-3016

Authorized officer

Maddox, A

INTERNATIONAL SEARCH REPORT

Internal Application No

PCT/EP 95/00263

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	NUCLEIC ACIDS RESEARCH, vol. 20, no. 1, 1992 OXFORD GB, pages 33-39, HIRLING, H., ET AL. 'Expression of active iron regulatory factor from a full-length human cDNA by in vitro transcription/translation' see figure 2 ---	4,5,10, 11
X	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 90, April 1993 WASHINGTON US, pages 3631-3635, YAMANASHI, Y., ET AL. 'Identification of HS1 protein as a major substrate of protein-tyrosine kinase(s) upon B-cell antigen receptor-mediated signaling' see figure 1c peptide A-2 ---	2
X	ELECTROPHORESIS 11, 528-536, 1990, BAUW, G., ET AL. 'Two-dimensional gel electrophoresis, protein electroblotting and microsequencing: a direct link between proteins and genes.' see page 534 sequence ief8318 fragment SAAEMYGSXFDLDYDFQ(R) ---	2
X	WO-A-93 21223 (UNIV OKLAHOMA STATE) 28 October 1993 see sequences MPPPGMRP and PIGMPPPG page 6 1.7,1.11 ---	2
X	WO-A-94 01451 (BIONEBRASKA INC) 20 January 1994 see sequence i.d. 9 see example 2 ---	2
P,X	WO-A-94 20622 (AKZO NOBEL NV ;KELDERMANS CORNELIA ELISABETH (NL); HORZINEK MARIAN) 15 September 1994 see sequence i.d. 7,fig.2,p.27 ---	2
P,X	EMBL SEQUENCE DATABASE ACCESSION NO.X82841 RELEASE 41; 28-11-1994. PEYRET, P., ET AL. 'A.THALINA ACO GENE' see sequence ---	1-11
P,X	EMBL SEQUENCE DATABASE ACCESSION NO.X82839 RELEASE 41,28-11-1994;PEYRET, P., ET AL. 'A.THALIANA MRNA FOR ACONITASE (ZAPII)' see sequence ---	1-6,9-11

	-/--	

INTERNATIONAL SEARCH REPORT

Internal Application No
PCT/EP 95/00263

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	EMBL SEQUENCE DATABASE ACCESSION NO.X82840 RELEASE 41, 28-11-1994, PEYRET, P., ET AL. 'C.MELO MRNA FOR ACONITASE (UNI-ZAPXR)' see sequence ---	1-6, 10, 11
A	PHYSIOLOGIA PLANTARUM, vol. 88, pages 485-492, DE BELLIS, L., ET AL. 'Purification and characterization of aconitase isoforms from etiolated pumpkin cotyledons' cited in the application see the whole document ---	1
A	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 88, WASHINGTON US, pages 10109-10113, KAPTAIN, S., ET AL. 'A regulated RNA binding protein also possesses aconitase activity' see the whole document ---	1-24
A	THE BIOCHEMICAL JOURNAL, vol. 276, pages 643-648, VERNIQUET, F., ET AL. 'Rapid inactivation of plant aconitase by hydrogen peroxide' cited in the application see the whole document ---	1-24
A	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 89, WASHINGTON US, pages 11730-11734, KENNEDY, M.C., ET AL. 'Purification and characterization of cytosolic aconitase from beef liver and its relationship to the iron-responsive element binding protein' see the whole document -----	1-24

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/EP 95/00263

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-9321223	28-10-93	AU-B- 4102893 EP-A- 0640101	18-11-93 01-03-95
WO-A-9401451	20-01-94	AU-B- 4772593 CA-A- 2140002 EP-A- 0651761	31-01-94 20-01-94 10-05-95
WO-A-9420622	15-09-94	NONE	

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale No
PCT/EP 95/00263

A. CLASSEMENT DE L'OBJET DE LA DEMANDE CIB 6 C12N15/60 C12N15/82 A01H5/00 C12P21/08		
Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB		
B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE Documentation minimale consultée (système de classification suivi des symboles de classement) CIB 6 C12N A01H C12P		
Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche		
Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si cela est réalisable, termes de recherche utilisés)		
C. DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	THE BIOCHEMICAL JOURNAL, vol. 294, pages 103-107, COURTOIS-VERNIQUET, F., ET AL. 'Lack of aconitase in glyoxysomes and peroxisomes' voir page 104, colonne de droite, ligne 3 - ligne 15 <div style="text-align: center;">---</div>	16
X	<div style="text-align: center;">---</div> PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 87, WASHINGTON US, pages 7958-7962, ROUAULT, T.R., ET AL. 'Cloning of the cDNA encoding an RNA regulatory protein - the human iron-responsive element binding-protein' cité dans la demande voir figure 2 <div style="text-align: center;">---</div>	4,5,10, 11
-/--		
<div style="display: flex; justify-content: space-between;"> <input checked="" type="checkbox"/> Voir la suite du cadre C pour la fin de la liste des documents <input checked="" type="checkbox"/> Les documents de familles de brevets sont indiqués en annexe </div>		
* Catégories spéciales de documents cités:		
"A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent "E" document antérieur, mais publié à la date de dépôt international ou après cette date "L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée) "O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens "P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée	"T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention "X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément "Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier "&" document qui fait partie de la même famille de brevets	
Date à laquelle la recherche internationale a été effectivement achevée <div style="text-align: center; font-weight: bold;">22 Juin 1995</div>		Date d'expédition du présent rapport de recherche internationale <div style="text-align: center; font-weight: bold;">- 3. 07. 95</div>
Nom et adresse postale de l'administration chargée de la recherche internationale Office Européen des Brevets, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tél. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016		Fonctionnaire autorisé <div style="text-align: center; font-weight: bold;">Maddox, A</div>

RAPPORT DE RECHERCHE INTERNATIONALE

 Demande internationale No
 PCT/EP 95/00263

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	NUCLEIC ACIDS RESEARCH, vol. 20, no. 1, 1992 OXFORD GB, pages 33-39, HIRLING, H., ET AL. 'Expression of active iron regulatory factor from a full-length human cDNA by in vitro transcription/translation' voir figure 2 ---	4,5,10, 11
X	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 90, Avril 1993 WASHINGTON US, pages 3631-3635, YAMANASHI, Y., ET AL. 'Identification of HS1 protein as a major substrate of protein-tyrosine kinase(s) upon B-cell antigen receptor-mediated signaling' voir figure 1c peptide A-2 ---	2
X	ELECTROPHORESIS 11, 528-536, 1990, BAUW, G., ET AL. 'Two-dimensional gel electrophoresis, protein electroblotting and microsequencing: a direct link between proteins and genes.' voir page 534 séquence ief8318 fragment SAAEMYGSXFDLDYDFQ(R) ---	2
X	WO-A-93 21223 (UNIV OKLAHOMA STATE) 28 Octobre 1993 voir séquences MPPPGMRP et PIGMPPPG page 6 1.7,1.11 ---	2
X	WO-A-94 01451 (BIONEBRASKA INC) 20 Janvier 1994 voir séquence i.d. 9 voir exemple 2 ---	2
P,X	WO-A-94 20622 (AKZO NOBEL NV ;KELDERMANS CORNELIA ELISABETH (NL); HORZINEK MARIAN) 15 Septembre 1994 voir séquence i.d. 7,fig.2,p.27 ---	2
P,X	EMBL SEQUENCE DATABASE ACCESSION NO.X82841 RELEASE 41; 28-11-1994. PEYRET, P., ET AL. 'A.THALINA ACO GENE' voir séquence ---	1-11
P,X	EMBL SEQUENCE DATABASE ACCESSION NO.X82839 RELEASE 41,28-11-1994;PEYRET, P., ET AL. 'A.THALIANA MRNA FOR ACONITASE (ZAPII)' voir séquence ---	1-6,9-11

	-/--	

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale No
PCT/EP 95/00263

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
P,X	EMBL SEQUENCE DATABASE ACCESSION NO.X82840 RELEASE 41, 28-11-1994, PEYRET, P., ET AL. 'C.MELO MRNA FOR ACONITASE (UNI-ZAPXR)' voir séquence ---	1-6, 10, 11
A	PHYSIOLOGIA PLANTARUM, vol. 88, pages 485-492, DE BELLIS, L., ET AL. 'Purification and characterization of aconitase isoforms from etiolated pumpkin cotyledons' cité dans la demande voir le document en entier ---	1
A	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 88, WASHINGTON US, pages 10109-10113, KAPTAIN, S., ET AL. 'A regulated RNA binding protein also possesses aconitase activity' voir le document en entier ---	1-24
A	THE BIOCHEMICAL JOURNAL, vol. 276, pages 643-648, VERNIQUET, F., ET AL. 'Rapid inactivation of plant aconitase by hydrogen peroxide' cité dans la demande voir le document en entier ---	1-24
A	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 89, WASHINGTON US, pages 11730-11734, KENNEDY, M.C., ET AL. 'Purification and characterization of cytosolic aconitase from beef liver and its relationship to the iron-responsive element binding protein' voir le document en entier -----	1-24

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demar: internationale No

PCT/EP 95/00263

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO-A-9321223	28-10-93	AU-B- 4102893 EP-A- 0640101	18-11-93 01-03-95
WO-A-9401451	20-01-94	AU-B- 4772593 CA-A- 2140002 EP-A- 0651761	31-01-94 20-01-94 10-05-95
WO-A-9420622	15-09-94	AUCUN	